BMJ Open

Development of machine learning support for reading whole body diffusion weighted magnetic resonance imaging (WB-MRI) in myeloma for the detection and quantification of the extent of disease before and after treatment (MALIMAR): protocol for a cross-sectional diagnostic test accuracy study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-067140
Article Type:	Protocol
Date Submitted by the Author:	03-Aug-2022
Complete List of Authors:	Satchwell, Laura; Royal Marsden Hospital NHS Trust Wedlake, Linda; Royal Marsden Hospital NHS Trust Greenlay, Emily; Royal Marsden Hospital NHS Trust Li, Xingfeng; Imperial College London, Department of Cancer and Surgery Messiou, Christina; Royal Marsden Hospital NHS Trust; Institute of Cancer Research Glocker, Ben ; Imperial College London Department of Computing Barwick, Tara; Imperial College London, Department of Cancer and Surgery; Imperial College Healthcare NHS Trust, Department of Radiology Barfoot, Theodore; King's College London Doran, Simon; Institute of Cancer Research Leach, Martin O; Institute of Cancer Research Koh, Dow Mu; Royal Marsden Hospital NHS Trust; Institute of Cancer Research Sutton Kaiser, Martin; Institute of Cancer Research; Royal Marsden Hospital NHS Trust Winzeck, Stefan; Imperial College London Department of Computing Qaiser, Talha; Imperial College London Department of Computing Aboagye, Eric; Imperial College London Department of Surgery and Cancer ROCKALL, ANDREA; Imperial College London Department of Surgery and Cancer; Imperial College Healthcare NHS Trust, Department of Radiology
Keywords:	Myeloma < HAEMATOLOGY, Diagnostic radiology < RADIOLOGY & IMAGING, Magnetic resonance imaging < RADIOLOGY & IMAGING, ONCOLOGY

Development of machine learning support for reading whole body diffusion weighted magnetic resonance imaging (WB-MRI) in myeloma for the detection and quantification of the extent of disease before and after treatment (MALIMAR): protocol for a cross-sectional diagnostic test accuracy study

Authors:

Laura C Potts[1]; Linda Wedlake[1]; Emily Greenlay[1]; Xingfeng Li[3]; Christina Messiou[1,6]; Ben Glocker[2]; Tara Barwick[3,4]; Theodore Barfoot[5]; Simon Doran[6]; Martin O Leach[6]; Dow Mu Koh[1,6]; Martin Kaiser[1,6]; Stefan Winzeck[2]; Talha Qaiser[2]; Eric O.Aboagye[3]; Andrea Rockall[3,4]

[1]Royal Marsden Hospital NHS Foundation Trust

[2] BioMedIA Group, Department of Computing, Imperial College London

[3]Department of Cancer and Surgery, Imperial College London

[4]Department of Radiology, Imperial College Healthcare NHS Trust

[5]Kings College London

[6]Institute of Cancer Research

Corresponding author: Miss Laura C Potts; laura.potts@rmh.nhs.uk. ORCID ID: 0000-0002-2935-6532

Abbreviations:

ADC: Apparent Diffusion Coefficient

BMI: Body Mass Index

BRC: Biomedical Research Centre

CCR: Committee for Clinical Research

CPMS: Central Portfolio Management System

CRF: Case Report Form

CRN: Clinical Research Network

CRUK: Cancer Research UK

CT: Computerised Tomography

CTIMP: Clinical Trial of Investigational Medicinal Product

DWI: Diffusion Weighted Imaging

EME: Efficacy and Mechanism Evaluation

F-FDG: F-Fluorodeoxyglucose

HRA: Health Research Authority

HV: Healthy Volunteer

ICHT: Imperial College Healthcare Trust

ICR: The Institute of Cancer Research

IRAS: Integrated Research Application System

ML: Machine Learning

MM: Multiple Myeloma

MRC: Medical Research Council

MRI: Magnetic Resonance Imaging

MS: Microsoft

NHS: National Health Service

NICE: National Institute of Clinical Excellence

NIHR: National Institute of Health Research

PET: Positron Emission Tomography

PPI: Patient and Public Involvement

QA: Quality Assurance

REC: Research Ethics Committee

sFLC: serum Free Light Chain

TMG: Trial Management Group

TSC: Trial Steering Committee

WB-DW-MRI: Whole Body Diffusion Weighted Magnetic Resonance Imaging

UKRI: UK Research and Innovation

Acknowledgements:

We acknowledge NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden and Institute of Cancer Research and the NIHR Royal Marsden Clinical Research Facility.

We acknowledge the support of the Imperial College London NIHR BRC Imaging Theme and the CRUK Imperial Centre and the Imaging Research Office at ICHT.

We acknowledge the support of the CRUK funded National Cancer Imaging Translational Accelerator award (Institute of Cancer Research and Imperial College London).

Availability of data and materials:

Not applicable.

Ethics approval and consent to participate:

The Royal Marsden NHS Foundation Trust is the study sponsor and responsible for initiating and managing the study, for oversight of the conduct of the study including submission of financial returns to NIHR, for reporting trial progress in line with NIHR EME requirements, for submitting and co-ordinating amendments to the trial protocol and for oversight of closure and archiving of all trial materials. All publications must have the consent of the NIHR. The study protocol was reviewed by the Royal Marsden NHS Foundation Trust and Institute of Cancer Research Combined Committee for Clinical Research (CCR) and underwent proportionate review by the South Central – Oxford C Research Ethics Committee in November 2017 (IRAS Project ID: 233501) and the Health Research Authority. The study was also approved for CPMS Portfolio adoption (CPMS ID: 36766). This research will be carried out in accordance with the Declaration of Helsinki (1996). The study will be conducted in accordance with the conditions of ethical approval. Participation is limited to Healthy Volunteers only. Before participation all participants will be provided with a Healthy Volunteer Participation Sheet and will give written informed consent.

Consent for publication:

Not applicable.

Protocol version: 3.0 31/01/2019

Abstract:

Introduction: Whole-body MRI (WB-MRI) is recommended by NICE as the first-line imaging tool for diagnosis of multiple myeloma. Reporting WB-MRI scans requires expertise to interpret and can be challenging for radiologists who need to meet rapid turn-around requirements. Automated computational tools based on machine learning (ML) could assist the radiologist in terms of sensitivity and reading speed and would facilitate improved accuracy, productivity and cost-effectiveness. The MALIMAR study aims to develop and validate a ML algorithm to increase the diagnostic accuracy and reading speed of radiological interpretation of WB-MRI compared to standard methods.

Methods and analysis: This phase II/III imaging trial will perform retrospective analysis of previously obtained clinical radiology MRI scans and scans from healthy volunteers obtained prospectively to implement training and validation of a machine learning algorithm. The study will comprise three project phases using approximately 633 scans to 1) train the ML algorithm to identify active disease; 2) clinically validate the ML algorithm; and 3) determine change in disease status following treatment via a quantification of burden of disease in myeloma patients. Phase 1 will primarily train the ML algorithm to detect active myeloma against an expert assessment ('reference standard'). Phase 2 will utilise the ML output in the setting of radiology reader study to assess the difference in sensitivity when using ML-assisted reading or human-alone reading. Phase 3 will assess the agreement between experienced readers (with and without ML) and the reference standard in scoring both overall burden of disease before and after treatment, and response.

Ethics and dissemination: MALIMAR has ethical approval from South Central – Oxford C Research Ethics Committee (REC Reference: 17/SC/0630). MALIMAR is funded by National Institute for Healthcare Research Efficacy and Mechanism Evaluation funding (NIHR EME Project ID: 16/68/34). Findings will be made available through peer-reviewed publications and conference dissemination.

Trial registration: The study was registered at clincaltrials.gov (NCT03574454) on 2 July 2018.

Strengths and limitations of this study:

- This cross-sectional diagnostic test accuracy study will be the first of its kind to provide evidence of whether a machine learning algorithm can deliver outputs to significantly enhance the radiology reading process for the benefit of myeloma patients.
- The MALIMAR study will explore whether specific processes can be automated to augment the ML process which could include the development of an automated segmentation tool to depict bony anatomy and volume of disease and negate the need for manual outlining.
- This study will provide ML outputs that can be tested across the NHS in live real-time clinical settings.
- This study will acquire and characterise what is possibly the largest set of myeloma WB-MRI scans in the UK; however, data will be acquired over a long period of time meaning scan quality could vary.
- Replicating clinical reporting in a retrospective study setting can be difficult to achieve, particularly for analysis of reading time.

Introduction

There is strong evidence in the existing literature for the use of whole-body MRI (WB-MRI) in the management of patients with multiple myeloma. In 2016, the National Institute of Clinical Excellence (NICE) made the recommendation of using WB-MRI as the first line imaging tool for diagnosis, based on the literature(1). A consensus from the International Myeloma Working Group agreed that identification of focal lesions more than 5mm on MRI should now be used as an indication to treat (2,3). Evidence suggests that diffusion weighted (DW) WB-MRI (WB-DW-MRI) is the most sensitive MR technique for detecting marrow disease (4–8) and superior to FDG-PET/CT for the detection of small sites of disease and diffuse infiltration (9,10). Therefore, WB-MRI is increasingly being adopted at centres worldwide for patients with myeloma. Treatment of high-risk patients is known to improve overall survival (11), therefore improved diagnostic accuracy is likely to translate into improved patient selection for treatment and prolonged survival.

Despite the acknowledged benefits of WB-MRI for patients with myeloma, with publication of the NICE guidance, one of the major concerns is how these complex scans can be reported by a radiology workforce in crisis. Specificity of disease detection in the marrow is improved by viewing source DW images alongside quantitative Apparent Diffusion Coefficient (ADC) maps. This allows differentiation of active sites of disease with restricted diffusion from treated sites of disease and vertebral haemangiomas which conversely return a very high ADC (12). Dixon images are also integral to image interpretation and morphological imaging is also necessary to identify mechanical complications of myeloma bone disease. Therefore, diagnostic accuracy is dependent on viewing multiple imaging sequences (7) and typically over 1200 image slices per WB-MRI scan in order to achieve whole body coverage. Consequently, reading time for the scans may be significant. At least 9% of UK radiology posts are unfilled (13) and in 2015 clinical radiology was placed on the national shortage occupation list. The time-consuming process of reporting WB-MRI scans is a concern for radiologists who need to provide rapid turn-around with a high productivity to support the NHS.

Automated computational tools based on machine learning (ML) could support reporting of these large datasets and facilitate translation of this valuable imaging technique into the NHS, not only in detecting active disease but also in identifying response to treatment. Ideally, a ML algorithm would automatically detect and highlight suspicious regions and could reduce reading time. An accurate and automatic detection of pathology may also increase diagnostic accuracy.

The possibility of using computer-assisted ML techniques has been considered in aiding interpretation of complex imaging datasets (14–16). Current work in the EME NIHR funded MALIBO study (17,18)(13/122/01) has demonstrated fully automatic multi-organ segmentation using WB-MRI in healthy volunteers (HV) and ML detection of primary colorectal cancer and metastatic lesions.

Aim

The aim of the MALIMAR study is to develop and validate a Machine Learning (ML) algorithm to improve the sensitivity of radiologists to detect the presence and extent of active myeloma before and after treatment, with high reproducibility and reduced reading time (WB-MRI with ML, the intervention) when compared with the standard of care radiology read (WB-MRI without ML support, the comparator).

Methods and analysis

Study design

The study is based on a cross-sectional diagnostic test accuracy design and will comprise three distinct project phases as summarised in Figure 1.

- In **Phase 1** the ML algorithm will be trained using both HV and myeloma patient scans to recognise active myeloma deposits as distinct from cases with no active disease, classifying disease as 'focal', 'diffuse' or 'inactive'.
- In **Phase 2** the ML algorithm will be validated using a second unseen dataset against a reference standard (i.e. ground truth) to assess how accurately radiologists classify disease using scans with the ML algorithm and compared to readings without ML. Diagnostic accuracy on a per patient and per region (using 16 pre-defined anatomical sites – Table 1) basis and reading time will be measured.
- In **Phase 3**, further development of the ML algorithm to quantify disease burden will be undertaken using datasets from phase 1 and 2. This quantification output will be tested in the phase 3 reader study in which readers will record disease burden and response between paired baseline (new diagnosis or relapse prior to initiation of treatment) and single post treatment WB-MRI scans, with or without ML support, and tested against the reference standard.

Table 1: Comparison of MALIMAR anatomical regions between ground truth CRFs and reader CRFs

Anatomical Regions	
Ground Truth CRFs (Phase 1 and 2)	**Reader CRFs (Phase 2)**
Skull	Skull
Scapula right	Ribs / clavicles / sternum / scapulae
Scapula left	Ribs / clavicles / sternum / scapulae
Clavicle right	Ribs / clavicles / sternum / scapulae
Clavicle left	Ribs / clavicles / sternum / scapulae
Sternum	Ribs / clavicles / sternum / scapulae
Spine upper	Cervical spine
Spine middle	Dorsal spine
Spine lower	Lumbar spine
Ribs right	Ribs / clavicles / sternum / scapulae

Ribs left	Ribs / clavicles / sternum / scapulae
Sacrum	Pelvis
Femur right	Long bones
Femur left	Long bones
Humerus right	Long bones
Humerus left	Long bones

Participants and Recruiting Centres

The study will be run at The Royal Marsden NHS Foundation Trust across two Royal Marsden Hospital (RMH) sites; Chelsea and Sutton, and Imperial College Healthcare Trust (ICHT). Patient and HV scans will make up the study population, and disease classification will be at both the scan and anatomical site level.

The scan population will comprise of; HV WB-MRI scans acquired from participants prospectively recruited from the Sponsor site only (RMH), with the option of the Imperial Site providing previously acquired HV scans; WB-MRI scans acquired as part of clinical care from patients being managed at RMH and ICHT; and WB-MRI scans previously acquired for a prospective research study in WB-MRI (iTIMM study). All scans acquired for the study will be done so using clinical standard of care Trust protocols.

The inclusion/exclusion criteria for the HV and patient scans are detailed in Table 2 and the planned number of scans for each study phase is detailed in Table 3.

Table 2: Inclusion and exclusion criteria.

	Inclusion criteria	Exclusion criteria
Healthy volunteers	Written informed consent No contra-indication to MRI 40 years or above in age (attempts will be made to include similar age range as myeloma patients) No known significant illness No known metallic implant	Significant artifact on scan Corrupted scan data
Patients in	Patient with confirmed myeloma with WB-MRI scan	Corrupted WB-MRI scan

phase 1 & 2	previously performed as part of clinical care. Sufficient imaging and clinical data for the expert reference panel to categorise the WB-MRI scan as: 1. Previously treated inactive disease with no evidence of active disease based on expert reference panel 2. Active disease – focal 3. Active disease – diffuse 4. Active disease – extra-medullary 5. New active myeloma, no previous treatment Patients may be included if the pattern of disease is a combination of focal, diffuse and/or extra-medullary.	data. Insufficient clinical data to allow the expert reference panel to categorise the scan.
Patients in phase 3	Training set: Phase 1 active disease cases and their post treatment scans from phase 2. Validation set: from iTIMM study. Written informed consent for iTIMM study All patients over the age of 18 with multiple myeloma planned for autograft.	Corrupted scan data. MRI incompatible metal implants Claustrophobia Diagnosis of other malignancy within 5 yrs

Table 3: Number of Healthy Volunteer (HV) and Multiple Myeloma (MM) scans in each category for each study phase.

	HV**	MM inactive	MM active focal	MM active diffuse	MM new diagnosis	Total
Phase 1*	40	40	60	40	20	200
Phase 2	50	100	105	70	28	353
Phase 3 training***	0	(80 post treatment)	60	40	20	200
Phase 3 validation	0	60 patients in iTIMM study scanned at baseline and post treatment				120

*The number of scans in phase 1 may increase by 140-180 scans (100 subjects) if there is evidence of over-fitting in the development of the algorithm.

** A total of 50 HV will be used, 40 in phase 1, which will be used again in phase 2, with the addition of 10 more HV.

***Scans used in phase 3 training are scans that have been previously used in phase 1 and 2

Intervention and Reference Standard

Intervention (including comparator)

The comparator in this study is defined as WB-MRI scans read by experienced radiologists, as per standard care (WB-MRI, the COMPARATOR). The intervention will use these standard methods with the addition of machine learning (WB-MRI+ML, the INTERVENTION). The ML algorithm will be

developed during phase 1 of the study following data curation and scan allocation to phase 1 and 2. DWI, ADC map and T1 weighted sequences (Dixon fat and water scans) will be used, reflecting the radiological reading tools used by expert readers.

Radiologists or readers are defined as experienced based on their previous clinical radiology reading skills and responsibilities, and their length of service in this role. Experienced readers will be required to have completed at least 100 WB-MRI clinical scan reports.

Reference standard

There is no available histological reference standard for every site of bone marrow disease, as trephine biopsy is usually restricted to a single site. The proposed reference standard thus comprises the interpretation of an expert panel; a radiologist and a haematologist who are experts in myeloma. They will have access to 1) WB-MR images; 2) bone marrow histopathology reports (with quantitation); 3) serum paraproteins; 4) serum free light chain (sFLC), in order to categorise per scan:

- Presence or absence of active disease
- The detailed disease distribution by anatomical site
- Quantitation of the burden of disease (using a validated MRI score (19,20) and sFLC) including category of response to treatment .

Scan and site level data from these scans will be captured on Case Report Forms (CRFs) for all cases in phases 1 and 2 and used as 'ground truth' in the classification of study output. Reference standard for phase 3 will be obtained from the source (iTIMM study).

Objectives

Primary research objectives

Phase 1: To develop a myeloma-specific ML algorithm to detect the presence of active disease on WB-MRI+ML (with machine learning '+ML') with sufficient sensitivity.

Phase 2: To validate WB-MRI+ML against the comparator WB-MRI for sensitivity on a per-patient and per site basis.

Phase 3: To develop and validate a ML algorithm to automatically quantify the burden of active disease, before and after treatment.

Secondary research objectives (Phase 2 and 3 only)

For each of the following, our objective is to compare WB-MRI with and without ML support to the reference standard for:

1) Reading time

2) Specificity

3) Sensitivity of non-experienced readers

4) Agreement of categorising disease as focal, diffuse and/or extramedullary.

5) Agreement of categorising patients as responder or non-responder

Procedure

Scan Acquisition – Healthy Volunteers

Healthy Volunteers (HV) will be recruited to obtain data from normal bone marrow within the age range typical of myeloma. Up to 50 HVs aged 40 years or above will be recruited using approved advertisements at the Sponsor site and consented with the help of CRN resources (See Supplementary S1a for consent form). The HV Information Sheet (Supplementary file S1b) will clearly explain the MRI scanning procedure and the actions that will be taken in the event of incidental (i.e. unexpected) findings. Contact details will be supplied on the HV Information Sheet to enable volunteers to respond to the invitation or ask any questions. A total of 22 HV scans previously acquired are also available for use from ICHT if needed.

Participating HVs will undergo a single whole body MRI scan at RMH according to the trial specific scanning protocol. HV scans will be acquired in the following sequences (T1, fat/water, Dixon, ADC, etc) to mirror the clinical setting and on Siemens, Avanto and Aero (wide bore) MRI scanners (Supplementary S2 details sequences). Subjects with a larger BMI will be scanned on the Siemens Aero which has a larger bore diameter to optimise comfort.

Scan Acquisition– Myeloma Patients

Previously acquired patient scans will be identified by the investigators within the Sponsor's myeloma clinical service (between 2011 and 2020), supplemented by scans from ICHT, until the required sample size is reached. Scans will normally include the following sequences; T1, fat/water, Dixon, ADC, etc, and on the following MRI machines; Siemens, Avanto and Aero MRI scanners (Supplementary S2 for sequence details).

Scan Classification and Allocation to Study Phase

Patient scans will be categorised by the expert reference panel as showing inactive disease, active focal, active diffuse (focal or diffuse) and new disease. HV scans will be classified as normal (i.e. non-diseased). Scans will be allocated to Phase 1 or 2 as per Table 3. To minimise bias or 'over-learning', no more than 5 scans from the same patient will be allocated to Phase 1. Phase 2 scans will not include any patient scans that have been used in Phase 1 and thus comprise only those previously unseen by the ML algorithm. A subset of scans from phase 1 and 2 will be used to further train the algorithm at the start of Phase 3. Phase 3 validation scans have previously been acquired for the iTTiM trial (NCT02403102) and include a unique series of paired scans, previously unseen by the ML algorithm.

Scan Curation (Quality Control) and Anatomical Segmentation

Eligible scans will be curated immediately prior to transfer to an online platform for secure storage (ICR XNAT). This will ensure the ML algorithm is able to interpret all scans consistently. Curation scripts will be written in python and ensure that scans exhibit consistent characteristics such as:

correct sequential display of images, no missing slices, noting presence of unusual artifacts that might interrupt ML reads and other factors which might compromise interpretation. Further details on the data curation will be published elsewhere.

Phase 1 scans will then be manually segmented into 16 bone regions (Table 1) using a boundary box approach. These scans will be used to teach the ML algorithm to recognise active myeloma disease (focal or diffuse) and precision metrics will be evaluated in order to achieve the optimal algorithm. Initially, scans will be classified by the ML algorithm at scan level (i.e. patient level) only.

Testing of ML Algorithm – Radiology Reading Process

The ML algorithm will be tested by both experienced and inexperienced radiology readers.

Phase 2 scans will be subjected to the ML algorithm which will provide an ML overlay on all scans indicating areas of disease by means of a heat map. For each scan, a 'standard' and 'machine learning' version will be available. The trial statistician will randomly allocate reads to each of the (approximately 15 – 20) readers, using trial-specific algorithms written using Stata software (StataCorp, Texas). The reads will be performed in two batches to incorporate a wash-out period. Each batch will have 50% of cases with ML support and 50% without, to avoid reader training bias. The reading process will be described in a Reader Manual and all readers will receive appropriate training in viewing scans using the Biotronics 3D Web-based platform and completing a Read CRF available via MS Forms (see supplementary file S3a). In the case of "inexperienced" readers, training will comprise a review of the CRFs and the viewing software with a basic training on reporting lexicon. A scribe will be provided to assist readers during the reading process and input data to the CRF in each batch of reads. Following a 4-week wash out period, readers will be presented with the second batch of reads with the opposite reading paradigm with regards to the ML support. The same cases will be allocated to the same readers. A subset of approximately 50 scans will be read a second time by a different reader as an interrater check.

In Phase 3, scans from the iTIMM study, comprising paired baseline and follow-up post treatment scans, will be used to test whether the ML algorithm is capable of distinguishing change in disease status (i.e. disease burden) between the two time-points. Reads will again be randomly allocated to the readers by the trial statistician. Readers will follow similar procedures to that outlined above with one set of paired scans having the ML overlay and the other with no ML overlay (for CRF see supplementary file S3b). A 4-week wash-out period will again apply between the two batches of reads. A subset of approximately 20 scans will be read a second time by a different reader as an interrater check.

Data collection

Reader responses will be captured using MS Forms with responses being transferred directly to an excel spreadsheet. Examples of the CRFs to be used in both ML validation phases are given as supplemental files (S3a, S3b). All readers will be provided with a manual describing CRF completion (including a lexicon of disease definitions) and use of the software viewing tools and overlay of the ML output heatmap and opportunity for live training using the online platform.

Outcome measures

Phase 1 – ML Algorithm Training Phase

Primary: Sensitivity for the detection of active myeloma on WB-MRI + ML detection tool against the reference standard.

Secondary: 1. Specificity; 2. F1 score (a single measure of precision and recall).

Phase 2 – ML Algorithm Clinical Testing Phase (Presence /Absence of active myeloma)

Primary: Difference in sensitivity of WB-MRI -/+ ML detection tool to diagnose the presence of active myeloma on a per-patient basis, by experienced readers, assessed against the reference standard.

Secondary: For comparison of WB-MRI -/+ML: 1. Per-site sensitivity to diagnose active disease; 2. Reading time; 3. Specificity; 4. Agreement with reference standard to categorise disease as focal, diffuse and/or extramedullary; 5. Sensitivity of non-experienced readers for presence of active disease.

Phase 3 – ML Algorithm for quantification of disease burden with clinical testing

Primary : agreement between experienced readers and the reference standard in scoring overall burden of disease before and after treatment for response categorisation -/+ ML quantification tool.

Secondary: For comparison of WB-MRI -/+ML: 1. Reading time; 2. Agreement of categorisation of patients as responder or non-responder with the reference standard; 3. Agreement of non-experienced readers for burden of disease and categorisation of response; 4. Estimated difference in cost for radiology reading time for WB-MRI -/+ML.

Proposed tertiary: Verification of the team's previously published work regarding reverse classification accuracy: predicting segmentation performance in the absence of a reference standard (21).

Sample size

Phase 1:

We will train the ML algorithm on a set of scans without and with active disease that will reflect the categories of disease that may be encountered in clinical practice. The number of cases used for training is arbitrarily chosen reflecting the knowledge that a large number of training datasets will improve training accuracy, counterbalanced with the resources needed to curate and annotate a large number of datasets.

Phase 2:

The study is powered on the primary outcome of sensitivity.

In a meta-analysis, Wu et al have reported a pooled sensitivity of 88% and a pooled specificity of 86% (0.86 for WB-MRI with DW-MRI) (8). We anticipate that the addition of ML could increase this by at least 7.5%, from 88% to 95.5%. There is no background data to indicate the expected proportion of discordant pairs so we have estimated this as (1-0.955)*0.88 + 0.955*(1-0.88), which is equal to 0.154. To achieve 80% power using a two-sided alpha of 0.05 would require a total of 203 patients positive for myeloma using the gold standard.

If it is assumed that the specificity will be unchanged using ML, a total number of cases with no active disease of 150 (50 HV, 100 inactive treated myeloma), will give 80% power to show that the difference is above a non-inferiority limit of 10%.

Phase 3 training:

Approximately 200 cases that have at least two time points will be taken from phase 1 and 2, with active disease present at least at one time point, and used for training and validation for burden of disease; this will ensure efficient use of all data and segmentations.

Phase 3 clinical testing:

This sample size is fixed at 60 patients, the full sample size of the iTIMM study, each of whom has a baseline and one post treatment scan.

Statistical Analysis

Phase 1 analysis

The ability to correctly localise and detect active disease will be evaluated by calculating sensitivity, specificity and the F1 score (a single measure of precision (positive predictive value) and recall (sensitivity)) for multiple algorithms and compared against the reference standard. Following TSC approval, the optimal algorithm will move forward to phase 2.

Phase 2 analysis

In phase 2, the percentage of patients with active disease on WB-MRI +/- ML support who have positive reference standard will be compared using McNemar's test with a two-sided alpha of 0.05. Per patient and per site sensitivity and specificity with and without ML support will be reported with 95% confidence intervals. Reading time will be compared using Wilcoxon's test for paired data and described using summary statistics.

The same analysis of sensitivity, specificity and reading time will be repeated for inexperienced readers.

Agreement between experienced and inexperienced readers will be measured in a subset of cases with a Kappa coefficient, and overall proportion of concordant cases.

All other endpoints will be summarised using descriptive statistics.

Although the study is powered to detect superiority of the primary endpoint, if sensitivity is shown to be non-inferior using ML and reading time is both clinically and statistically significantly lower using ML, this would be considered as an indication to proceed. Non-inferiority in this context will be defined as having any possible reduction in sensitivity with ML significantly higher than a lower limit of -10% (using Tangos' test with one-sided alpha 0.05)

Phase 3 analysis

In phase 3, the difference between the experienced readers' disease score to the reference standard disease score will be recorded and compared +/- ML support using Wilcoxon's test. Differences from scores given by experienced readers and the reference standard will be described using Bland-Altman plots for scores +/- ML support.

All other endpoints will be summarised using descriptive statistics.

A simple cost-effectiveness analysis may be performed depending on study findings, such as the reading time.

Procedure(s) to account for missing or spurious data

If a scan is incomplete or the file is corrupted and not evaluable, it will be excluded from the dataset. If a set of radiology reads is incomplete, a new trained reader will be identified to do the full allocation of reads.

Timing and responsibility for analyses

Analyses will take place at both the end of phase 2 and then again at the end of phase 3, when all readings have been completed.

Patient and public involvement (PPI)

A PPI representative was appointed from an established group at Myeloma UK. The individual gave in-depth feedback on the study, particularly on the relevance to patient care and the use of retrospective patient data and HV scans. Myeloma UK is fully supportive of the project and is willing to assist with dissemination of important findings to the Myeloma UK community.

Safety

As this study is recruiting healthy volunteers only, an a-priori agreement has been reached with the Sponsor that safety reporting is not required. Sponsor procedures in respect of incidental (i.e. unexpected) findings in healthy volunteers will be adhered to and results captured within the Trust's Clinical Record.

Monitoring against Source Data will not be required which is in line with the Sponsor's policy on non-CTIMP trials.

Trial funding, organisation and administration

The study has been awarded funding by MRC NIHR EME (Efficacy and Mechanism Evaluation) Awards Body (NIHR EME Project ID: 16/68/34). In addition, the Department of Radiology has agreed to fund the cost of HV WB-MRI scans. The cost of recruitment and consenting of HVs will be requested through the NHS Clinical Research Network. RMH is the study sponsor responsible for

initiating and managing the study and the coordinating centre, including sign-off of the study protocol.

A Trial Management Group (TMG) meeting will be held regularly to ensure satisfactory progress of the study. A Trial Steering Committee (TSC) will provide independent oversight for the study, review the development of the ML algorithm, and advise the TMG where problems may arise. The TSC will include a Patient Advocate.

Ethics and dissemination

Ethical approval for MALIMAR was granted on 21/11/2017 (REC) and 21/12/2017 (HRA) Here, we report version 3.0 of the protocol. All participating sites gained local approval prior to study participation.

Any protocol modifications will be submitted for approval to the REC, reflected in the online registration and disseminated by e-mail to site principal investigators and trial coordinators. The statistician will have access to the final linked trial dataset. There are no plans to provide public access to the full protocol, participant-level data, or statistical code. The researchers aim to publish results in a peer-reviewed journal and share via social media and conferences. Authorship will be determined according to academic standards.

Discussion

This study aims to develop and validate a ML algorithm to augment the performance and efficiency of the radiology reading process using WB-MRI. The results will show the impact of using the ML tool and outcomes of the study will have implications for the application of ML with WB-MRI in myeloma patients across the NHS. It is anticipated that feasibility analysis will follow the successful completion of this study to pilot the implementation of the ML tool in a real-time prospective study prior to future clinical setting.

To avoid bias we ensure: 1) comparator and intervention tests are read by readers that are fully blinded to the reference standard; 2) a mixture of cases with and without disease; 3) the reads will be presented such that radiologists must read a mixture of cases without or with ML support during each round of reading including a wash-out period. We will have unavoidable incorporation bias, as the expert reference panel will use the MRI as part of the reference standard. The reference panel will consist of a single person's opinion which is a limitation to our study. If resources had allowed, the gold standard would have been to have two blinded opinions with a consensus panel in cases of disagreement. Other limitations include varying scan quality as data is acquired over a 9-year period; and replicating clinical reporting in a retrospective study setting can be challenging.

In conducting this study, we will have acquired possibly the largest set of characterised myeloma patient MRI scans in the UK and we anticipate this will form the basis of a unique training resource in the future.

Machine learning techniques in WB-MRI scans of patients with myeloma is likely to be transferable to other malignancies. In prostate and breast cancer, quantification of metastatic bone disease is an unmet need as bone only disease is not uncommon and is currently classified as non-measurable by RECIST 1.1 (22). The participating HVs will be consented to allow the anonymised datasets to be a future resource for the wider research community.

Study status

The MALIMAR study opened on 26 April 2018 using protocol version 1.0 (30 Oct 2017). The study was in phase II, using protocol version 3.0 (31 Jan 2019), at date of submission. Protocol amendments are documented in Supplementary S4.

Author contributions:

AR, CM, TaB, BG, SW, TQ, ThB, SD, ML, MK and DK conceptualization and methodology; AR, CM, TaB, ThB, MK, BG, TQ, XF and SW investigation; EA and AR resources; ThB and SD data curation; LP and EG formal analysis; AR, DK and CM supervision; LP writing – original draft; AR,

CM, TaB, BG, LW and LP writing – review and editing; BG, TQ, XF and SW data visualisation; LW project administration; EA and AR funding acquisition.

Competing interests:

AR receives honoraria for educational lecture at Garmisch International Symposium, has an unpaid role on the European Society of Radiology Board of Directors and receives travel cost support where necessary. BG receives grants from other entities; EU commission and UKRI London Medical Imaging & Artificial Intelligence Centre for Value Based Healthcare, is a Scientific advisor for Kheiron Medical Technologies (Jan 2018 – Sep 2021) and receives stock options as part of standard employment packages from both Kheiron Medical Technologies and HeartFlow. EA has a patent pending for Machine Learning in Alzheimer's disease and has a role on the scientific advisory board for Radiopharm Theranostics Limited. MK receives grants from both Myeloma UK and Celgene/BMS, and consulting fees or payments from AbbVie, BMS/Celgene, Janssen, GSK, Karyopharm, Takeda and Seagen. CM receives additional funding as a co-investigator on a radiology NIHR study and is part of the joint venture Celescan with the Royal Marsden, The Institute of Cancer Research and Sopra Steria. TB receives additional funding from CRUK grant funding (NCITA) and NIHR (HTA) and receives honoraria from Bayer.

Funding:

This study (ID: 16/68/34) is funded by the Efficacy and Mechanism Evaluation (EME) Programme, an MRC and NIHR partnership. In addition, the Department of Radiology has agreed to fund the cost of healthy volunteer whole body MRI scans. The cost of recruitment and consenting of healthy volunteers will be requested through the NHS Clinical Research Network. The views expressed in this publication are those of the authors and not necessarily those of the MRC, NHS, the NIHR, or the Department of Health and Social Care.

EG and LP's posts are part funded by the National Institute for Health and Care Research (NIHR) Biomedical Research Centre at The Royal Marsden NHS Foundation Trust and the Institute of Cancer

Research, London. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

SW is supported by the UKRI London Medical Imaging & Artificial Intelligence Centre for Value Based Healthcare.

Data sharing statement:

Anonymised data are available upon reasonable request as a resource for the wider research community from Andrea Rockall (ORCID ID: 0000-0001-8270-5597), providing consent has been granted from all participants.

References

1. NICE. Myeloma: diagnosis and management NICE guideline [NG35]. 2016.
2. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. The lancet oncology. 2014;15(12):e538–48.
3. Dimopoulos MA, Hillengass J, Usmani S, Zamagni E, Lentzsch S, Davies FE, et al. Role of magnetic resonance imaging in the management of patients with multiple myeloma: a consensus statement. 2015;
4. Pearce T, Philip S, Brown J, Koh DM, Burn PR. Bone metastases from prostate, breast and multiple myeloma: differences in lesion conspicuity at short-tau inversion recovery and diffusion-weighted MRI. Br J Radiol. 2012;85(1016):1102–6.
5. Squillaci E, Manenti G, di Stefano F, Miano R, Strigari L, Simonetti G. Diffusion-weighted MR imaging in the evaluation of renal tumours. Journal of Experimental and Clinical Cancer Research. 2004;23(1):39–46.
6. Dutoit JC, Vanderkerken MA, Anthonissen J, Dochy F, Verstraete KL. The diagnostic value of SE MRI and DWI of the spine in patients with monoclonal gammopathy of undetermined significance, smouldering myeloma and multiple myeloma. Eur Radiol. 2014;24(11):2754–65.
7. Messiou C, Hillengass J, Delorme S, Lecouvet FE, Moulopoulos LA, Collins DJ, et al. Guidelines for acquisition, interpretation, and reporting of whole-body MRI in myeloma: myeloma response assessment and diagnosis system (MY-RADS). Radiology. 2019;291(1):5–13.
8. Wu L, Gu H, Zheng J, Xu X, Lin L, Deng X, et al. Diagnostic value of whole-body magnetic resonance imaging for bone metastases: a systematic review and meta-analysis. Journal of Magnetic Resonance Imaging. 2011;34(1):128–35.
9. Messiou C, Porta N, Sharma B, Levine D, Koh DM, Boyd K, et al. Prospective evaluation of whole-body MRI versus FDG PET/CT for lesion detection in participants with myeloma. Radiology: Imaging Cancer. 2021;3(5).

10. Pawlyn C, Fowkes L, Otero S, Jones JR, Boyd KD, Davies FE, et al. Whole-body diffusion-weighted MRI: a new gold standard for assessing disease burden in patients with multiple myeloma? Leukemia. 2016;30(6):1446–8.

11. Mateos MV, Hernández MT, Giraldo P, de la Rubia J, de Arriba F, Corral LL, et al. Lenalidomide plus Dexamethasone for High-Risk Smoldering Multiple Myeloma. New England Journal of Medicine. 2013 Aug;369(5):438–47.

12. Padhani AR, Koh DM, Collins DJ. Whole-body diffusion-weighted MR imaging in cancer: current status and research directions. Radiology. 2011;261(3):700–18.

13. Radiologists RC of. Clinical radiology UK workforce census 2015 report. The Royal College of Radiologists London; 2016.

14. Juntu J, Sijbers J, de Backer S, Rajan J, van Dyck D. Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine. 2010;31(3):680–9.

15. Pauly O, Glocker B, Criminisi A, Mateus D, Möller AM, Nekolla S, et al. Fast multiple organ detection and localization in whole-body MR Dixon sequences. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2011. p. 239–47.

16. Lavdas I, Rockall AG, Castelli F, Sandhu RS, Papadaki A, Honeyfield L, et al. Apparent diffusion coefficient of normal abdominal organs and bone marrow from whole-body DWI at 1.5 T: the effect of sex and age. American Journal of Roentgenology. 2015;205(2):242–50.

17. Lavdas I, Glocker B, Rueckert D, Taylor SA, Aboagye EO, Rockall AG. Machine learning in whole-body MRI: experiences and challenges from an applied study using multicentre data. Clin Radiol. 2019;74(5):346–56.

18. Lavdas I, Glocker B, Kamnitsas K, Rueckert D, Mair H, Sandhu A, et al. Fully automatic, multiorgan segmentation in normal whole body magnetic resonance imaging (MRI), using classification forests (CF s), convolutional neural networks (CNN s), and a multi-atlas (MA) approach. Med Phys. 2017;44(10):5210–20.

19. Giles SL, Desouza NM, Collins DJ, Morgan VA, West S, Davies FE, et al. Assessing myeloma bone disease with whole-body diffusion-weighted imaging: comparison with x-ray skeletal survey by region and relationship with laboratory estimates of disease burden. Clin Radiol. 2015;70(6):614–21.

20. Giles SL, Messiou C, Collins DJ, Morgan VA, Simpkin CJ, West S, et al. Whole-body diffusion-weighted MR imaging for assessment of treatment response in myeloma. Radiology. 2014;271(3):785–94.

21. Valindria V v, Lavdas I, Bai W, Kamnitsas K, Aboagye EO, Rockall AG, et al. Reverse classification accuracy: predicting segmentation performance in the absence of ground truth. IEEE Trans Med Imaging. 2017;36(8):1597–606.

22. Lecouvet FE, Talbot JN, Messiou C, Bourguet P, Liu Y, de Souza NM, et al. Monitoring the response of bone metastases to treatment with Magnetic Resonance Imaging and nuclear medicine techniques: a review and position statement by the European Organisation for Research and Treatment of Cancer imaging group. Eur J Cancer. 2014;50(15):2519–31.

576

577 **Figure 1: MALIMAR Study Flow Diagram**

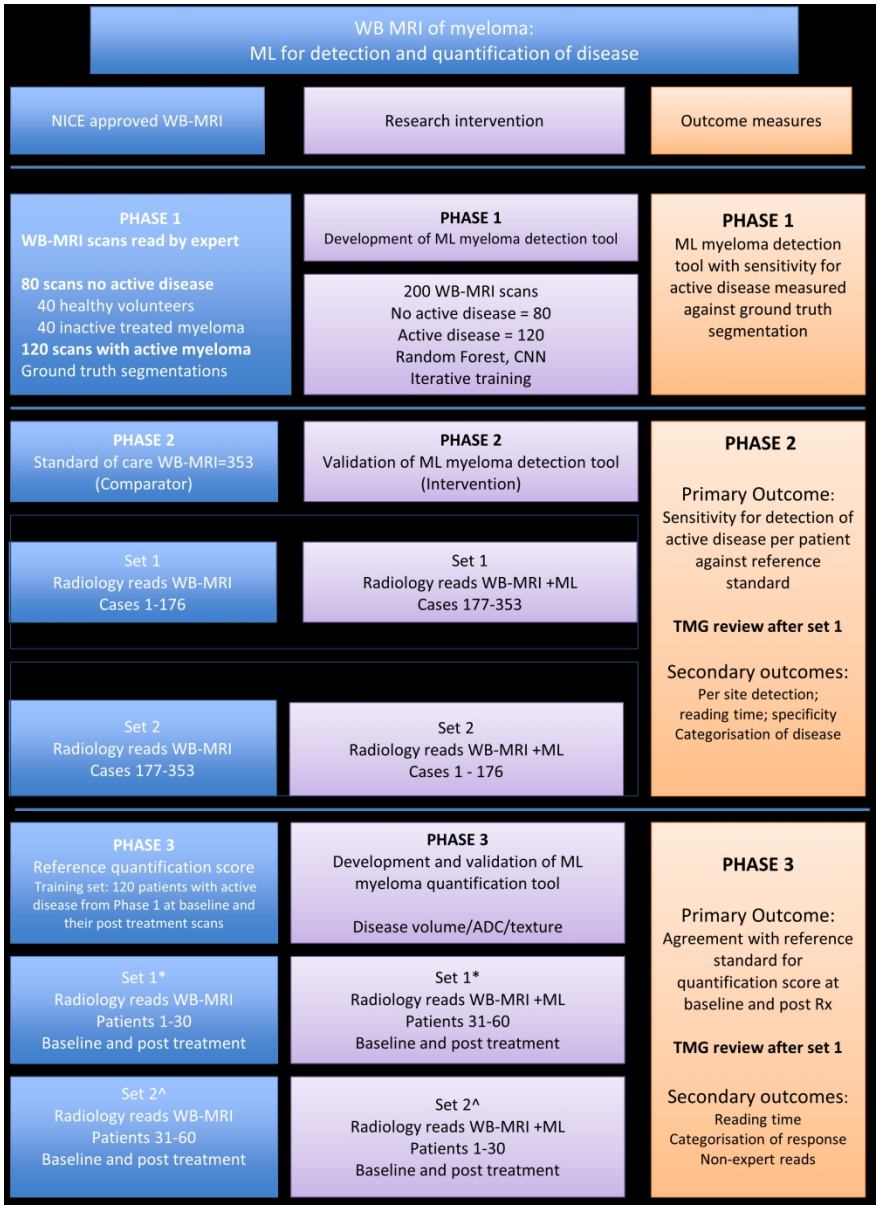

Figure 1: Study Flow Chart

526x723mm (130 x 130 DPI)

The ROYAL MARSDEN
NHS Foundation Trust

The MALIMAR Study Healthy Volunteer Consent Form

Study Reference Numbers: CCR 4820: IRAS No.: 233501

NHS No. ☐☐☐☐☐☐☐☐☐☐

Healthy volunteer Trial ID: ☐☐☐☐☐☐☐☐☐

Name of Lead Researcher: ……

Please initial box

1. I confirm that I have read and understand the Healthy Volunteer Information Sheet version 2.0 dated 07/12/18 for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐
3. If I request withdrawal from the study, I give permission that my data already collected within the study can be anonymised and used. ☐
4. I understand that relevant sections of my medical notes may be looked at by responsible individuals from the research team, from regulatory authorities or from the NHS trust where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
5. I consent to undergo an MRI scan under the supervision of the responsible clinician for this research. I understand that if any health related issues come to light as a result of undergoing this scan, otherwise known as 'incidental findings', that I and my General Practitioner will be promptly informed of these issues. ☐
6. I agree to participate in the MALIMAR study. ☐
7. I give permission for the data collected during the study to be used in further ethically approved research within and outside the UK in the field of imaging research. I understand this will not include any personal data from which I could be identified. ☐

Name of Healthy Volunteer	Date	Signature
Name of person taking consent (PI or approved signatory)	Date	Signature

Original for Investigator's Site File; 1 copy for volunteer; 1 copy for hospital notes; 1 copy to be sent to RM-CTU

The ROYAL MARSDEN
NHS Foundation Trust

MALIMAR

Healthy Volunteer Information Sheet

Development of machine learning support for reading whole body diffusion weighted magnetic resonance imaging (WB-DW-MRI) in myeloma for the detection and quantification of the extent of disease before and after treatment.

Short Title: MAchine **L**earning **I**n **M**yelom**A** **R**esponse

7th December 2017

Version 2.0

CCR Number: 4820

IRAS (Integrated Research Application System) No. 233501

You are being invited to take part in a research study. Before you decide whether or not to take part it is important for you to understand why we are doing this research and what it involves. Please take time to read the following information carefully and discuss it with relatives, friends, and your GP if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time deciding whether or not you wish to take part.

You can learn more about clinical research on the Cancer Research UK's patient website (www.cancerhelp.org.uk)

Invitation

If you are 40 years or above the Radiology Department at the Royal Marsden hospital would like to invite you to take part in a research study. This will involve you having a particular type of Magnetic Resonance Imaging (MRI) scan known as a Whole-Body Diffusion Weighted MRI scan or 'WB-DW-MRI'.

Before you decide to participate it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with friends, relatives and your GP if you wish. Ask us if there is anything that is not clear or if you would like more information. Thank you for reading this Information Sheet.

What is the purpose of the study?

There have been enormous advances in recent years in the technology used to take pictures (images) of the internal anatomy of cancer patients to better identify sites of disease. These images (or scans) can now provide a more accurate indication of the scope or spread of disease. They can also be used for assessing disease response to different drugs or treatments.

MRI (magnetic resonance imaging) has the advantage over other types of scanning (e.g. computerised tomography or 'CT') in that it does not involve the delivery of any radiation dose. In particular, a new type of MRI, called Whole Body Diffusion Weighted MRI (WB-DW-MRI) can provide especially precise images of diseased compared to healthy tissues. As a result, it is now being more widely used in cancer treatment centers throughout the world.

Despite these advantages, WB-DW-MRI has an important disadvantage. Each scan is made up from over a thousand images, each of which needs to be read and interpreted by an expert Radiologist. Thus, the time taken to read a single WB-DW-MRI scan is much longer than for a normal MRI scan, meaning that few NHS treatment centres (or hospitals) are able to offer them to patients.

Members of the research team from the Royal Marsden Hospital and Imperial College London have already undertaken some work to ascertain how computers can reduce the time taken to read WB-DW-MRI scans. The technique is called 'machine learning' and basically teaches a computer to detect areas of suspicion or concern for disease on WB-DW-MRI scans. The 'trained' computer can then make an initial and very rapid interpretation of the images taken during a scan. These images can then be presented to the expert radiologist to make the final interpretation. In addition to training computers to read scans more quickly, we also want to train computers to interpret differences between scans taken from the same patient at different time-points. This will allow us to accurately assess change in disease extent or response to treatment over time.

However, in order to train the computers, we need examples of WB-DW-MRI images taken from both diseased (cancerous) and healthy tissues. In this study we are concentrating on patients with myeloma (cancer of white blood cells). We have already acquired WB-DW-MRI images from many patients with this type of cancer. **So now, we are seeking your help to acquire WB-DW-MRI images from healthy tissues for the M**achine **L**earning **I**n **M**yelom**A R**esponse (**MALIMAR**) study**.**

What will happen to me if I decide that I would like to take part?

Before we can enter you to the study, we will need to check that you can have an MRI scan and that you are suitable to take part. Some people cannot have an MRI scan. These include people with a pacemaker, metal heart valves, aneurysm clips in the brain or people who have had metal fragments in their eyes. In addition, we are unable to include volunteers who have had or have a significant illness as this may affect the scan.

It may also not be appropriate for you to take part if you have had extensive surgery previously. Our study researcher will confirm these points with you before you are admitted to the trial. As advised above we are only recruiting volunteers aged 40 and above: anyone under this age will have to be

excluded from participating because they will not be a suitable comparator. Once we have confirmed that you are suitable to enter the trial, we will ask you to sign an Informed Consent Form and then book your scan. Some volunteers may be asked to attend early evening or week-end appointments to avoid busy times during the day when the MRI Unit is reserved for patients. There are usually no special preparations and no injection or drugs will be given. All instructions for the scan will be in your MRI appointment letter. When you come for the scan you are advised to wear clothing without metal fastenings and to avoid using make-up or mascara. You can wear glasses, but will need to take these off during the scan. A locker will be provided for your valuables.

The MRI scan will be carried out by radiographers who are trained to carry out the scans. MRI uses a magnetic field and radio waves to build up detailed images of your internal anatomy by detecting signals sent out by water molecules. It is not painful, but you will have to lie still for the duration of the scan which can be up to 60 minutes. The scanner produces a variety of loud noises during the scan which are made by the magnetic coils that switch on and off during the scan. These are important in measuring the signals from your body to create the images. They are switched on and off very quickly and they vibrate, which is what causes the noise.

Some people may find the noise level uncomfortable and the table quite hard to lie on. You will be provided with earplugs to help reduce the noise. The scanner is open at both ends, but some people may find it claustrophobic. During the scan the radiographer can see you from the control room and can talk to you through an intercom. You will be given a call button to press to alert attention and can listen to music during the scan. You can leave as soon as your scan is finished and can eat and drink as normal. There are no side effects from the MRI scan itself.

Why am I being invited to take part?

You will be reading this Information Sheet because you have responded to one of our advertisements for Healthy Volunteers to take part. If we invite you to sign a Consent Form then you are eligible to take part in the study. If you are not eligible to participate we will explain the reason.

Do I have to take part?

No, it is up to you to decide whether or not to take part. If you do choose to take part you will be asked to sign a consent form, a copy of which will be given to you for your records along with this information sheet about the study. Your legal rights are not affected by participation in the study.

What happens if I change my mind during the study?

Your participation in this study is entirely voluntary. If you agree to take part and then change your mind and wish to withdraw, you may do so at any time. If you decide to not join the study or to discontinue in the study, this will not affect any future care or treatment you receive.

What are the risks and the benefits of taking part in this study?

A possible risk in taking part is a degree of discomfort you may encounter in undergoing the MRI scan. As we said above, unlike other forms of imaging (e.g. CT scans) MRI does not deliver radiation and no drugs or other medication will be given. You will be registered on the Royal Marsden Hospital Information System and a report of your scan results will be held on this system. If an unexpected finding of concern is discovered, a doctor will call you to discuss your scan report. We will also send a copy of the report to your GP who will then advise you regarding any follow-up investigations that may be needed. This could lead to some anxiety. If unexpected findings are discovered which are not concerning, we will send you a letter to explain the findings and copy this letter to your GP. You may then wish to call us or your GP for more information. If there are no unexpected findings we will not contact you or your GP.

In general, the research will not be of direct benefit to you, but may prove to be of benefit to others in the future. However, possible benefits are that you may find it satisfying to have contributed to medical research and, should an unexpected finding be discovered you may feel that the early detection and diagnosis will result in a better outcome. If you wish to have a copy of your scan report, you may ask for this.

What if something goes wrong?

It is unlikely that anything will go wrong but, if you wish to complain, you can do so using the normal NHS complaints procedure. If taking part harms you in any way, there are no special compensation arrangements, but the hospital would be liable for any negligence on the part of hospital staff. Your legal rights are not affected by giving your consent to participate in this study.

Who is organizing and funding the research?

This study is being organised by The Royal Marsden NHS Foundation Trust with participation from The Institute of Cancer Research, Imperial College London and Imperial Healthcare NHS Foundation Trust. The study is being funded by a National Institute for Health Research grant as part of their Efficacy and Mechanism Evaluation programme.

Will my taking part in this study be kept confidential?

1) Clinical Information: You will need to be given a Royal Marsden hospital number in order to receive the WB-DW-MRI scan. The resulting scan report will be held on our clinical Hospital Information (NHS PACS) System which is the system we use for holding all NHS patient information. Access to this system is subject to the normal Trust-based information governance controls. If, in the event of unexpected findings, you require further diagnostic investigations, your GP will be informed and your scans and accompanying data will be made available to the hospital treating you.

2) Research Information: Your scan data will be anonymised and identified by a unique trial identification number. Your unique trial number will be used to make sure you cannot be identified by members of the research team that are not part of the NHS staff at RMH. The data from your scan which will be used in the MALIMAR study will only be available to authorised members of our research team so they can collect information needed for this research study and also to check that it is correct. All information will be kept confidential, and your name, date of birth and other identifiable information will be removed from your scans prior to archiving. We will also ask you to consent to allow your data that has been collected in the study to be sent outside of the UK and to be used in future ethically approved studies. This information will not include any personal information that could directly identify you.

What will happen to the results of this study?

As soon as there are reliable results, they will be published in a respected peer reviewed medical journal and presented in various scientific meetings. Your identity will not be revealed in any report, publication or presentation. The results will be available on request.

How is the trial monitored for safety?

This study has been carefully planned by leading cancer specialists and approved by the Oxford C Research Ethics Committee (REC), the Royal Marsden Hospital Committee for Clinical Research (CCR) and the Health Research Authority (HRA). The members of the study team will be meeting at regular intervals to monitor the progress and safety of the study. Full (100%) monitoring will be carried out to ensure that where incidental findings come to light, both you and your GP are promptly informed.

What do I do now?

We would be happy to answer any questions you may have about the study. You can telephone us, or speak to us again. Please discuss this information with your family, friends or your GP if you wish. If you require further information about this study please contact:

Professor Andrea Rockall,
Chief Investigator,
Clinical Chair Radiology,
ICTEM Building,
Imperial College Healthcare HNS Trust,
Du Cane Road
London, W12 0NN
Tel: 0207 59 42792 (Personal Assistant to Professor Rockall)

Dr Christina Messiou,
Principal Investigator,
Consultant Radiologist,
The Royal Marsden NHS Foundation Trust,
Fulham Road
London, SW3 6JJ
Tel: 0208 661 3216

Veronica Morgan
MRI Research Superintendent Radiographer
Clinical Magnetic Resonance Unit, Sutton
The Royal Marsden NHS Foundation Trust
Tel 02089156493

Thank you for reading and considering taking part in this study.

Funding Acknowledgement: Funding from the National Institute for Health Research – Efficacy and Mechanism Evaluation (NIHR – EME) programme for the MALIMAR study is acknowledged.

Participant Type	Study name	Site	MRI Machine Name	Sequences acquired
Healthy Volunteers	MALIMAR	Royal Marsden	Siemens Aera	Haste localiser, Sag T1, T2 whole spine, DWI (B50,600,900 and ADC), fl3d_CAIPI_wb_tra_BH_20 and T2 HASTE Vertex to knees
Healthy Volunteers	MALIMAR	Royal Marsden	Siemens Avanto	localiser, Sag T1, T2 whole spine, DWI (B50,600,900 and ADC), fl3d_vibe_dixon_TRA_15deg 256_pocS and T2 HASTE Vertex to knees
Myeloma Patients	MALIMAR	Royal Marsden	Siemens Aera	Haste localiser, Sag T1, T2 whole spine, DWI (B50,600,900 and ADC), fl3d_CAIPI_wb_tra_BH_20 and Vertex to knees
Myeloma Patients	MALIMAR	Royal Marsden	Siemens Avanto	localiser, Sag T1, T2 whole spine, DWI (B50,600,900 and fl3d_vibe_dixon_TRA_15deg 256 Vertex to knees
Myeloma Patients	MALIMAR	ICHT	Siemens Aera	Axial dixons x 4 350 slices each (total: 1400) B 50 248 slices B900 248 slices ADC 248 slices Sag T1 spine 15 slices Sag T2 spine 15 slices

MALIMAR Radiology Reads - CRF Phase 2

Version 4, 06 September 2021

* Required

1. Scan ID *

2. Reader ID *

3. Round *

- ◯ Round 1
- ◯ Round 2

4. Date of Read *

Please input date (dd/MM/yyyy)

5. Start time of read - Enter in format: HH:MM using 24 hour clock *

6. Disease status - BONES - Record Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Cervical Spine	◯	◯	◯	◯
Dorsal Spine	◯	◯	◯	◯
Lumbar Spine	◯	◯	◯	◯
Pelvis	◯	◯	◯	◯
Long Bones	◯	◯	◯	◯
Skull	◯	◯	◯	◯
Ribs / Clavicles / Sternum / Scapulae	◯	◯	◯	◯

7. Disease status - BONES - Record maximum size of Active / Focal lesions (mm) *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Cervical Spine	◯	◯	◯	◯
Dorsal spine	◯	◯	◯	◯
Lumbar spine	◯	◯	◯	◯
Pelvis	◯	◯	◯	◯
Long Bones (max. long axis)	◯	◯	◯	◯
Skull	◯	◯	◯	◯
Ribs / Clavicles / Sternum / Scapulae (max. long axis)	◯	◯	◯	◯

8. Disease Status - BONES - How confident are you in your assessment of Active / Focal lesions *

	Not at all confident	Some confidence	Confident	Very Confident
Cervical Spine	○	○	○	○
Dorsal Spine	○	○	○	○
Lumbar Spine	○	○	○	○
Pelvis	○	○	○	○
Long Bones	○	○	○	○
Skull	○	○	○	○
Ribs / Clavicles / Sternum / Scapulae	○	○	○	○

9. Disease Status - Record if diffuse disease was present at any of these sites? *

	Yes	No
Cervical Spine	◯	◯
Dorsal Spine	◯	◯
Lumbar Spine	◯	◯
Pelvis	◯	◯
Long Bones	◯	◯
Skull	◯	◯
Ribs / Clavicles / Sternum / Scapulae	◯	◯

10. How confident were you in your assessment of diffuse disease at these sites? *

	Not at all confident	Some confidence	Confident	Very confident
Cervical Spine	◯	◯	◯	◯
Dorsal Spine	◯	◯	◯	◯
Lumbar Spine	◯	◯	◯	◯
Pelvis	◯	◯	◯	◯
Long Bones	◯	◯	◯	◯
Skull	◯	◯	◯	◯
Ribs / Clavicles / Sternum / Scapulae	◯	◯	◯	◯

11. Was extramedullary disease present at any site? *

◯ Yes

◯ No

12. If extramedullary disease was present at any site - state location(s) separated by a semi-colon

13. If extramedullary disease was present, what was your level of confidence in assessing this? *

- ◯ Not confident at all
- ◯ Some confidence
- ◯ Confident
- ◯ Very confident
- ◯ Not Applicable, no extramedullary disease is seen.

14. Confidence in assessing overall disease status on this scan (i.e. in determining the presence or absence of ANY active disease) *

- ◯ Not confident at all
- ◯ Some confidence
- ◯ Confident
- ◯ Very confident

15. Stop time of read - RECORD IMMEDIATELY AFTER COMPLETING CLINICAL READ - Enter in format: HH:MM using 24 hour clock *

16. TO BE COMPLETED FOLLOWING THE CLINICAL READ:
Was a Machine Learning Image available *

- ◯ Yes
- ◯ No

17. If a Machine Learning 'ML' Image was available, please indicate whether sites were positive for active / focal disease, i.e. was there an ML finding?

	Highly likely negative on ML	Probably negative on ML	Probably positive on ML	Highly likely positive on ML
Cervical Spine	◯	◯	◯	◯
Dorsal Spine	◯	◯	◯	◯
Lumbar Spine	◯	◯	◯	◯
Pelvis	◯	◯	◯	◯
Long Bones	◯	◯	◯	◯
Skull	◯	◯	◯	◯
Ribs / Clavicles / Sternum / Scapulae	◯	◯	◯	◯

18. If a Machine Learning 'ML' image was available, please indicate whether sites were positive for diffuse disease, i.e. was there an ML finding?

	Highly likely negative on ML	Probably negative on ML	Probably positve on ML	Highly likely positive on ML
Cervical Spine	○	○	○	○
Dorsal Spine	○	○	○	○
Lumbar Spine	○	○	○	○
Pelvis	○	○	○	○
Long Bones	○	○	○	○
Skull	○	○	○	○
Ribs / Clavicles / Sternum / Scapulae	○	○	○	○

19. Scan Quality: What was the quality of the WB-MRI used for this read? *

	Good	Adequate	Poor
1. B 900	○	○	○
2. ADC	○	○	○
3. T1 sequences	○	○	○

20. Please enter any specific comments you have on scan quality

21. Reader confirmation: My responses have been accurately reported on this CRF (enter 'yes' if in agreement with this statement) *

- ◯ Yes
- ◯ No

MALIMAR Radiology Reads - CRF Phase 3

Version 2, 31 March 2022

* Required

1. Scan ID Post Treatment Scan (PT) *

2. Scan ID - Baseline Scan (BL) *

3. Reader ID *

4. Phase 3 - Round *

- ◯ Round 1
- ◯ Round 2

5. Date of Read *

Format: M/d/yyyy

6. Start time of read - Enter in format: HH:MM using 24 hour clock *

7. CERVICAL SPINE - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

8. CERVICAL SPINE - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

9. CERVICAL SPINE - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

10. DORSAL SPINE - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

11. DORSAL SPINE - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

12. DORSAL SPINE - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

13. LUMBAR SPINE - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

14. LUMBAR SPINE - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

15. LUMBAR SPINE - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

16. PELVIS - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

17. PELVIS - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

18. PELVIS - Was Diffuse Disease present? *

	Yes	No
Post Treatment	◯	◯
Baseline	◯	◯

19. LONG BONES - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	◯	◯	◯	◯
Baseline	◯	◯	◯	◯

20. LONG BONES - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	◯	◯	◯	◯
Baseline	◯	◯	◯	◯

21. LONG BONES - Was Diffuse Disease present? *

	Yes	No
Post Treatment	◯	◯
Baseline	◯	◯

22. SKULL - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

23. SKULL - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

24. SKULL - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

25. RIBS / CLAVICLES / STERNUM / SCAPULAE - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

26. RIBS / CLAVICLES / STERNUM / SCAPULAE - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

27. RIBS / CLAVICLES / STERNUM / SCAPULAE - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

28. Was extramedullary disease present at any site? *

○ Yes

○ No

29. If extramedullary disease was present at any site - state location(s) separated by a semi-colon

30. If extramedullary disease was present, what was your level of confidence in assessing this? *

- ◯ Not confident at all
- ◯ Some confidence
- ◯ Confident
- ◯ Very confident
- ◯ Not Applicable, no extramedullary disease is seen.

31. OVERALL RESPONSE - Change in Disease Status (Baseline - Post-Treatment) *

	Complete Response	Partial Response	Stable Disease	Disease Progression
Response category	◯	◯	◯	◯

32. OVERALL RESPONSE - CONFIDENCE - How confident were you in assessing overall response? *

	Not at all confident	Some confidence	Confident	Very confident
Confidence category	◯	◯	◯	◯

33. Stop time of read - RECORD IMMEDIATELY AFTER COMPLETING CLINICAL READ - Enter in format: HH:MM using 24 hour clock *

34. TO BE COMPLETED FOLLOWING THE CLINICAL READ:
Was a Machine Learning Image available *

- ◯ Yes
- ◯ No

35. If Machine Learning 'ML' Images were available, please indicate category of response suggested by ML

	Complete Response	Partial Response	Stable Disease	Progressive Disease
Response category	◯	◯	◯	◯

36. Scan Quality: What was the quality of the WB-MRI used for this read? *

	Good	Adequate	Poor
1. B 900	◯	◯	◯
2. ADC	◯	◯	◯
3. T1 sequences	◯	◯	◯

37. Please enter any specific comments you have on scan quality

38. Reader confirmation: My responses have been accurately reported on this CRF (enter 'yes' if in agreement with this statement) *

- ○ Yes
- ○ No

Supplementary S4 – MALIMAR Amendments

No. and Type of Amendment	Date approved	Brief Details of amendment
1. Non-substantial	25/06/2018	Protocol administrative updates
2. Non-substantial	15/01/2019	Communications to HVs
3. Non-substantial	19/03/2019	Update on scan numbers for protocol
4. Non-substantial	16/10/2019	Addition of ICHT site
5. Non-substantial	28/06/2019	Extension of project time-line and uplift in costs

SPIRIT Checklist for *Trials*

Complete this checklist by entering the page and line numbers where each of the items listed below can be found in your manuscript.

Your manuscript may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please state "n/a" and provide a short explanation. **Leaving an item blank or stating "n/a" without an explanation will lead to your manuscript being returned before review.**

Upload your completed checklist as an additional file when you submit to *Trials*. You must reference this additional file in the main text of your protocol submission. The completed SPIRIT figure must be included within the main body of the protocol text and can be downloaded here: http://www.spirit-statement.org/schedule-of-enrolment-interventions-and-assessments/

In your methods section, please state that you used the SPIRIT reporting guidelines, and cite them as:

Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin J, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. SPIRIT 2013 Explanation and Elaboration: Guidance for protocols of clinical trials. BMJ. 2013;346:e7586

	Reporting Item		Page and Line Number	Reason if not applicable
Administrative information				
Title	#1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	Page 1, line 1	
Trial registration	#2a	Trial identifier and registry name. If not yet registered, name of intended registry	Page 7, line 141	
Trial registration: data set	#2b	All items from the World Health Organization Trial Registration Data Set	n/a	Not a RCT
Protocol version	#3	Date and version identifier	Page 22, line 488	

Funding	#4	Sources and types of financial, material, and other support	Page 4, line 71	
Roles and responsibilities: contributorship	#5a	Names, affiliations, and roles of protocol contributors	Page 4, line 65	
Roles and responsibilities: sponsor contact information	#5b	Name and contact information for the trial sponsor	Page 4, line 86	
Roles and responsibilities: sponsor and funder	#5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	Page 4, line 86 Page 20, line 452	
Roles and responsibilities: committees	#5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	Page 20, line 440	
Introduction				
Background and rationale	#6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	Page 8, line 156	

Background and rationale: choice of comparators	#6b	Explanation for choice of comparators	Page 9, line 181 - 191	
Objectives	#7	Specific objectives or hypotheses	Page 9, line 191	
Trial design	#8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)	Page 9, line 199 Page 12 line 254	
Methods: Participants, interventions, and outcomes				
Study setting	#9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	Page 10, line 216	
Eligibility criteria	#10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	Page 10-11, line 221-228 and Table 1	
Interventions: description	#11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	Page 11, line 229	
Interventions: modifications	#11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving / worsening disease)	n/a	No modifications

Interventions: adherance	#11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return; laboratory tests)	n/a	No strategies or monitoring of adherance
Interventions: concomitant care	#11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	n/a	Use of scans only
Outcomes	#12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	Page 15, line 343	
Participant timeline	#13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Figure 1 and Table 3	Schedule and assessment of scans rather than participants
Sample size	#14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	Page 16, lines 366	
Recruitment	#15	Strategies for achieving adequate participant enrolment to reach target sample size	Page 12, line 271 Page 13, line 285	

Methods: Assignment of interventions (for controlled trials)				
Allocation: sequence generation	#16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	Page 10, lines 194-210	
Allocation concealment mechanism	#16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	Page 10, lines 200-210	
Allocation: implementation	#16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	Page 14, line 316 Page 15, line 327	
Blinding (masking)	#17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	Page 14, line 318 Page 21, line 470	
Blinding (masking): emergency unblinding	#17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	n/a	No blinding of intervention
Methods: Data collection, management, and analysis				

Data collection plan	#18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	Page 15, line 337	
Data collection plan: retention	#18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	n/a	No participant retention/follow-up
Data management	#19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	Page 15, line 340	
Statistics: outcomes	#20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	Page 17, line 391-428	
Statistics: additional analyses	#20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	n/a	Primary and secondary analysis included only

Statistics: analysis population and missing data	#20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	Page 19, line 421	
Methods: Monitoring				
Data monitoring: formal committee	#21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	Page 19 line 433 – line 450	
Data monitoring: interim analysis	#21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	Page18, line 396 Page 18, line 408	
Harms	#22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	Page 19, line 433	No adverse event reporting as not an interventional study on participants. Specified in manuscript.
Auditing	#23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	Page 19, line 438	
Ethics and dissemination				

Research ethics approval	#24	Plans for seeking research ethics committee / institutional review board (REC / IRB) approval	Page 20, line 452	
Protocol amendments	#25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC / IRBs, trial participants, trial registries, journals, regulators)	Page 20, line 456	
Consent or assent	#26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	Page 13, lines 272-279	
Consent or assent: ancillary studies	#26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	n/a	No ancillary studies
Confidentiality	#27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	Page 20, line 458-461 Page 22, line 484	
Declaration of interests	#28	Financial and other competing interests for principal investigators for the overall trial and each study site	Page 5, line 102	
Data access	#29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	Page 20, line 457-458	

Ancillary and post trial care	#30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	n/a	No participant level intervention
Dissemination policy: trial results	#31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	Page 20, lines 457-461	
Dissemination policy: authorship	#31b	Authorship eligibility guidelines and any intended use of professional writers	Page 20, lines 457-461	
Dissemination policy: reproducible research	#31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	Page 20, lines 457-461	
Appendices				
Informed consent materials	#32	Model consent form and other related documentation given to participants and authorised surrogates	Supplementary material S1a, S1b	
Biological specimens	#33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	n/a	No biological specimens collected

BMJ Open

Development of machine learning support for reading whole body diffusion weighted magnetic resonance imaging (WB-MRI) in myeloma for the detection and quantification of the extent of disease before and after treatment (MALIMAR): protocol for a cross-sectional diagnostic test accuracy study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-067140.R1
Article Type:	Protocol
Date Submitted by the Author:	18-Aug-2022
Complete List of Authors:	Satchwell, Laura; Royal Marsden Hospital NHS Trust Wedlake, Linda; Royal Marsden Hospital NHS Trust Greenlay, Emily; Royal Marsden Hospital NHS Trust Li, Xingfeng; Imperial College London, Department of Cancer and Surgery Messiou, Christina; Royal Marsden Hospital NHS Trust; Institute of Cancer Research Glocker, Ben ; Imperial College London Department of Computing Barwick, Tara; Imperial College London, Department of Cancer and Surgery; Imperial College Healthcare NHS Trust, Department of Radiology Barfoot, Theodore; King's College London Doran, Simon; Institute of Cancer Research Leach, Martin O; Institute of Cancer Research Koh, Dow Mu; Royal Marsden Hospital NHS Trust; Institute of Cancer Research Sutton Kaiser, Martin; Institute of Cancer Research; Royal Marsden Hospital NHS Trust Winzeck, Stefan; Imperial College London Department of Computing Qaiser, Talha; Imperial College London Department of Computing Aboagye, Eric; Imperial College London Department of Surgery and Cancer ROCKALL, ANDREA; Imperial College London Department of Surgery and Cancer; Imperial College Healthcare NHS Trust, Department of Radiology
Primary Subject Heading:	Radiology and imaging
Secondary Subject Heading:	Oncology
Keywords:	Myeloma < HAEMATOLOGY, Diagnostic radiology < RADIOLOGY & IMAGING, Magnetic resonance imaging < RADIOLOGY & IMAGING, ONCOLOGY

Development of machine learning support for reading whole body diffusion weighted magnetic resonance imaging (WB-MRI) in myeloma for the detection and quantification of the extent of disease before and after treatment (MALIMAR): protocol for a cross-sectional diagnostic test accuracy study

Authors:

Laura Satchwell[1]; Linda Wedlake[1]; Emily Greenlay[1]; Xingfeng Li[3]; Christina Messiou[1,6]; Ben Glocker[2]; Tara Barwick[3,4]; Theodore Barfoot[5]; Simon Doran[6]; Martin O Leach[6]; Dow Mu Koh[1,6]; Martin Kaiser[1,6]; Stefan Winzeck[2]; Talha Qaiser[2]; Eric O.Aboagye[3]; Andrea Rockall[3,4]

[1]Royal Marsden Hospital NHS Foundation Trust

[2] BioMedIA Group, Department of Computing, Imperial College London

[3]Department of Cancer and Surgery, Imperial College London

[4]Department of Radiology, Imperial College Healthcare NHS Trust

[5]Kings College London

[6]Institute of Cancer Research

Corresponding author: Laura Satchwell; laura.satchwell@rmh.nhs.uk. ORCID ID: 0000-0002-2935-6532

Abbreviations:

ADC: Apparent Diffusion Coefficient

BRC: Biomedical Research Centre

CCR: Committee for Clinical Research

CPMS: Central Portfolio Management System

CRF: Case Report Form

CRN: Clinical Research Network

CRUK: Cancer Research UK

CT: Computerised Tomography

CTIMP: Clinical Trial of Investigational Medicinal Product

DWI: Diffusion Weighted Imaging

EME: Efficacy and Mechanism Evaluation

FDG: Fluorodeoxyglucose

HV: Healthy Volunteer

ICHT: Imperial College Healthcare Trust

ICR: The Institute of Cancer Research

ML: Machine Learning

MM: Multiple Myeloma

MRC: Medical Research Council

MRI: Magnetic Resonance Imaging

NHS: National Health Service

NICE: National Institute of Clinical Excellence

NIHR: National Institute of Health Research

PET: Positron Emission Tomography

PPI: Patient and Public Involvement

REC: Research Ethics Committee

sFLC: serum Free Light Chain

TMG: Trial Management Group

TSC: Trial Steering Committee

WB-DW-MRI: Whole Body Diffusion Weighted Magnetic Resonance Imaging

UKRI: UK Research and Innovation

Acknowledgements:

We acknowledge NHS funding to the NIHR Biomedical Research Centre (BRC) at The Royal Marsden and Institute of Cancer Research and the NIHR Royal Marsden Clinical Research Facility.

We acknowledge the support of the Imperial College London NIHR BRC Imaging Theme and the Cancer Research UK (CRUK) Imperial Centre and the Imaging Research Office at ICHT.

We acknowledge the support of the CRUK funded National Cancer Imaging Translational Accelerator award (Institute of Cancer Research and Imperial College London).

Availability of data and materials:

Not applicable.

Ethics approval and consent to participate:

The Royal Marsden NHS Foundation Trust is the study sponsor and responsible for initiating and managing the study, for oversight of the conduct of the study including submission of financial returns to NIHR, for reporting trial progress in line with NIHR EME requirements, for submitting and co-ordinating amendments to the trial protocol and for oversight of closure and archiving of all trial materials. All publications must have the consent of the NIHR. The study protocol was reviewed by the Royal Marsden NHS Foundation Trust and Institute of Cancer Research Combined Committee for Clinical Research (CCR) and underwent proportionate review by the South Central – Oxford C Research Ethics Committee in November 2017 (Integrated Research Application System Project ID: 233501) and the Health Research Authority. The study was also approved for Central Portfolio Management System (CPMS) Portfolio adoption (CPMS ID: 36766). This research will be carried out in accordance with the Declaration of Helsinki (1996). The study will be conducted in accordance with the conditions of ethical approval. Participation is limited to Healthy Volunteers only. Before participation all participants will be provided with a Healthy Volunteer Participation Sheet and will give written informed consent.

Consent for publication:

Not applicable.

Protocol version: 3.0 31/01/2019

Abstract:

Introduction: Whole-body MRI (WB-MRI) is recommended by the National Institute of Clinical Excellence (NICE) as the first-line imaging tool for diagnosis of multiple myeloma. Reporting WB-MRI scans requires expertise to interpret and can be challenging for radiologists who need to meet rapid turn-around requirements. Automated computational tools based on machine learning (ML) could assist the radiologist in terms of sensitivity and reading speed and would facilitate improved accuracy, productivity and cost-effectiveness. The MALIMAR study aims to develop and validate a ML algorithm to increase the diagnostic accuracy and reading speed of radiological interpretation of WB-MRI compared to standard methods.

Methods and analysis: This phase II/III imaging trial will perform retrospective analysis of previously obtained clinical radiology magnetic resonance imaging (MRI) scans and scans from healthy volunteers obtained prospectively to implement training and validation of a machine learning algorithm. The study will comprise three project phases using approximately 633 scans to 1) train the ML algorithm to identify active disease; 2) clinically validate the ML algorithm; and 3) determine change in disease status following treatment via a quantification of burden of disease in myeloma patients. Phase 1 will primarily train the ML algorithm to detect active myeloma against an expert assessment ('reference standard'). Phase 2 will utilise the ML output in the setting of radiology reader study to assess the difference in sensitivity when using ML-assisted reading or human-alone reading. Phase 3 will assess the agreement between experienced readers (with and without ML) and the reference standard in scoring both overall burden of disease before and after treatment, and response.

Ethics and dissemination: MALIMAR has ethical approval from South Central – Oxford C Research Ethics Committee (REC Reference: 17/SC/0630). MALIMAR is funded by National Institute for Healthcare Research Efficacy and Mechanism Evaluation funding (NIHR EME Project ID: 16/68/34). Findings will be made available through peer-reviewed publications and conference dissemination.

Trial registration: The study was registered at clincaltrials.gov (NCT03574454) on 2 July 2018.

Strengths and limitations of this study:

- The MALIMAR study has the potential to acquire and characterise what is possibly the largest set of myeloma WB-MRI scans in the UK.
- The cross-sectional diagnostic accuracy design allows for retrospective analysis of previously obtained clinical radiology scans for training and validation of a ML algorithm.
- This study will provide ML outputs that can be tested across the National Health Service in live real-time clinical settings.
- As data will be acquired over a long period of time, scan quality could vary.
- Replicating clinical reporting in a retrospective study setting can be difficult to achieve, particularly for analysis of scan reading time.

Introduction

There is strong evidence in the existing literature for the use of whole-body magnetic resonance imaging (WB-MRI) in the management of patients with multiple myeloma. In 2016, the National Institute of Clinical Excellence (NICE) made the recommendation of using WB-MRI as the first line imaging tool for diagnosis, based on the literature(1). A consensus from the International Myeloma Working Group agreed that identification of focal lesions more than 5mm on MRI should now be used as an indication to treat (2,3). Evidence suggests that diffusion weighted (DW) WB-MRI (WB-DW-MRI) is the most sensitive magnetic resonance technique for detecting marrow disease (4–8) and superior to Fluorodeoxyglucose Positron Emission Tomography / Computerised Tomography (FDG-PET/CT) for the detection of small sites of disease and diffuse infiltration (9,10). Therefore, WB-MRI is increasingly being adopted at centres worldwide for patients with myeloma. Treatment of high-risk patients is known to improve overall survival (11), therefore improved diagnostic accuracy is likely to translate into improved patient selection for treatment and prolonged survival.

Despite the acknowledged benefits of WB-MRI for patients with myeloma, with publication of the NICE guidance, one of the major concerns is how these complex scans can be reported by a radiology workforce in crisis. Specificity of disease detection in the marrow is improved by viewing source DW images alongside quantitative Apparent Diffusion Coefficient (ADC) maps. This allows differentiation of active sites of disease with restricted diffusion from treated sites of disease and vertebral haemangiomas which conversely return a very high ADC (12). Dixon images are also integral to image interpretation and morphological imaging is also necessary to identify mechanical complications of myeloma bone disease. Therefore, diagnostic accuracy is dependent on viewing multiple imaging sequences (7) and typically over 1200 image slices per WB-MRI scan in order to achieve whole body coverage. Consequently, reading time for the scans may be significant. At least 9% of UK radiology posts are unfilled (13) and in 2015 clinical radiology was placed on the national shortage occupation list. The time-consuming process of reporting WB-MRI scans is a concern for

radiologists who need to provide rapid turn-around with a high productivity to support the National Health Service (NHS). Automated computational tools based on machine learning (ML) could support reporting of these large datasets and facilitate translation of this valuable imaging technique into the NHS, not only in detecting active disease but also in identifying response to treatment. Ideally, a ML algorithm would automatically detect and highlight suspicious regions and could reduce reading time. An accurate and automatic detection of pathology may also increase diagnostic accuracy.

The possibility of using computer-assisted ML techniques has been considered in aiding interpretation of complex imaging datasets (14–16). Current work in the EME NIHR (Efficacy and Mechanism Evaluation National Institute of Health Research) funded MALIBO study (17,18)(13/122/01) has demonstrated fully automatic multi-organ segmentation using WB-MRI in healthy volunteers (HV) and ML detection of primary colorectal cancer and metastatic lesions.

Aim

The aim of the MALIMAR study is to develop and validate a Machine Learning (ML) algorithm to improve the sensitivity of radiologists to detect the presence and extent of active myeloma before and after treatment, with high reproducibility and reduced reading time (WB-MRI with ML, the intervention) when compared with the standard of care radiology read (WB-MRI without ML support, the comparator).

Methods and analysis

Study design

The study is based on a cross-sectional diagnostic test accuracy design and will comprise three distinct project phases as summarised in Figure 1.

- In **Phase 1** the ML algorithm will be trained using both HV and myeloma patient scans to recognise active myeloma deposits as distinct from cases with no active disease, classifying disease as 'focal', 'diffuse' or 'inactive'.
- In **Phase 2** the ML algorithm will be validated using a second unseen dataset against a reference standard (i.e. ground truth) to assess how accurately radiologists classify disease using scans with the ML algorithm and compared to readings without ML. Diagnostic accuracy on a per patient and per region (using 16 pre-defined anatomical sites – Table 1) basis and reading time will be measured.
- In **Phase 3**, further development of the ML algorithm to quantify disease burden will be undertaken using datasets from phase 1 and 2. This quantification output will be tested in the phase 3 reader study in which readers will record disease burden and response between paired baseline (new diagnosis or relapse prior to initiation of treatment) and single post treatment WB-MRI scans, with or without ML support, and tested against the reference standard.

Table 1: Comparison of MALIMAR anatomical regions between ground truth CRFs and reader CRFs

Anatomical Regions	
Ground Truth CRFs (Phase 1 and 2)	**Reader CRFs (Phase 2)**
Skull	Skull
Scapula right	Ribs / clavicles / sternum / scapulae
Scapula left	Ribs / clavicles / sternum / scapulae
Clavicle right	Ribs / clavicles / sternum / scapulae
Clavicle left	Ribs / clavicles / sternum / scapulae
Sternum	Ribs / clavicles / sternum / scapulae
Spine upper	Cervical spine
Spine middle	Dorsal spine
Spine lower	Lumbar spine
Ribs right	Ribs / clavicles / sternum / scapulae

Ribs left	Ribs / clavicles / sternum / scapulae
Sacrum	Pelvis
Femur right	Long bones
Femur left	Long bones
Humerus right	Long bones
Humerus left	Long bones

Participants and Recruiting Centres

The study will be run at The Royal Marsden NHS Foundation Trust across two Royal Marsden Hospital (RMH) sites; Chelsea and Sutton, and Imperial College Healthcare Trust (ICHT). Patient and HV scans will make up the study population, and disease classification will be at both the scan and anatomical site level.

The scan population will comprise of; HV WB-MRI scans acquired from participants prospectively recruited from the Sponsor site only (RMH), with the option of the Imperial Site providing previously acquired HV scans; WB-MRI scans acquired as part of clinical care from patients being managed at RMH and ICHT; and WB-MRI scans previously acquired for a prospective research study in WB-MRI (iTIMM study). All scans acquired for the study will be done so using clinical standard of care Trust protocols.

The inclusion/exclusion criteria for the HV and patient scans are detailed in Table 2 and the planned number of scans for each study phase is detailed in Table 3.

Table 2: Inclusion and exclusion criteria.

	Inclusion criteria	Exclusion criteria
Healthy volunteers	Written informed consent No contra-indication to MRI 40 years or above in age (attempts will be made to include similar age range as myeloma patients) No known significant illness No known metallic implant	Significant artifact on scan Corrupted scan data
Patients in	Patient with confirmed myeloma with WB-MRI scan	Corrupted WB-MRI scan

phase 1 & 2	previously performed as part of clinical care. Sufficient imaging and clinical data for the expert reference panel to categorise the WB-MRI scan as: 1. Previously treated inactive disease with no evidence of active disease based on expert reference panel 2. Active disease – focal 3. Active disease – diffuse 4. Active disease – extra-medullary 5. New active myeloma, no previous treatment Patients may be included if the pattern of disease is a combination of focal, diffuse and/or extra-medullary.	data. Insufficient clinical data to allow the expert reference panel to categorise the scan.
Patients in phase 3	Training set: Phase 1 active disease cases and their post treatment scans from phase 2. Validation set: from iTIMM study. Written informed consent for iTIMM study All patients over the age of 18 with multiple myeloma planned for autograft.	Corrupted scan data. MRI incompatible metal implants Claustrophobia Diagnosis of other malignancy within 5 yrs

Table 3: Number of Healthy Volunteer (HV) and Multiple Myeloma (MM) scans in each category for each study phase.

	HV**	MM inactive	MM active focal	MM active diffuse	MM new diagnosis	Total
Phase 1*	40	40	60	40	20	200
Phase 2	50	100	105	70	28	353
Phase 3 training***	0	(80 post treatment)	60	40	20	200
Phase 3 validation	0	60 patients in iTIMM study scanned at baseline and post treatment				120

*The number of scans in phase 1 may increase by 140-180 scans (100 subjects) if there is evidence of over-fitting in the development of the algorithm.

** A total of 50 HV will be used, 40 in phase 1, which will be used again in phase 2, with the addition of 10 more HV.

***Scans used in phase 3 training are scans that have been previously used in phase 1 and 2

Intervention and Reference Standard

Intervention (including comparator)

The comparator in this study is defined as WB-MRI scans read by experienced radiologists, as per standard care (WB-MRI, the COMPARATOR). The intervention will use these standard methods with the addition of machine learning (WB-MRI+ML, the INTERVENTION). The ML algorithm will be

developed during phase 1 of the study following data curation and scan allocation to phase 1 and 2. DW imaging (DWI), ADC map and T1 weighted sequences (Dixon fat and water scans) will be used, reflecting the radiological reading tools used by expert readers.

Radiologists or readers are defined as experienced based on their previous clinical radiology reading skills and responsibilities, and their length of service in this role. Experienced readers will be required to have completed at least 100 WB-MRI clinical scan reports.

Reference standard

There is no available histological reference standard for every site of bone marrow disease, as trephine biopsy is usually restricted to a single site. The proposed reference standard thus comprises the interpretation of an expert panel; a radiologist and a haematologist who are experts in myeloma. They will have access to 1) WB-MR images; 2) bone marrow histopathology reports (with quantitation); 3) serum paraproteins; 4) serum free light chain (sFLC), in order to categorise per scan:

- Presence or absence of active disease
- The detailed disease distribution by anatomical site
- Quantitation of the burden of disease (using a validated MRI score (19,20) and sFLC) including category of response to treatment .

Scan and site level data from these scans will be captured on Case Report Forms (CRFs) for all cases in phases 1 and 2 and used as 'ground truth' in the classification of study output. Reference standard for phase 3 will be obtained from the source (iTIMM study).

Objectives

Primary research objectives

Phase 1: To develop a myeloma-specific ML algorithm to detect the presence of active disease on WB-MRI+ML (with machine learning '+ML') with sufficient sensitivity.

Phase 2: To validate WB-MRI+ML against the comparator WB-MRI for sensitivity on a per-patient and per site basis.

Phase 3: To develop and validate a ML algorithm to automatically quantify the burden of active disease, before and after treatment.

Secondary research objectives (Phase 2 and 3 only)

For each of the following, our objective is to compare WB-MRI with and without ML support to the reference standard for:

1) Reading time

2) Specificity

3) Sensitivity of non-experienced readers

4) Agreement of categorising disease as focal, diffuse and/or extramedullary.

5) Agreement of categorising patients as responder or non-responder

Procedure

Scan Acquisition – Healthy Volunteers

Healthy Volunteers (HV) will be recruited to obtain data from normal bone marrow within the age range typical of myeloma. Up to 50 HVs aged 40 years or above will be recruited using approved advertisements at the Sponsor site and consented with the help of Clinical Research Network (CRN) resources (See Supplementary S1a for consent form). The HV Information Sheet (Supplementary file S1b) will clearly explain the MRI scanning procedure and the actions that will be taken in the event of incidental (i.e. unexpected) findings. Contact details will be supplied on the HV Information Sheet to enable volunteers to respond to the invitation or ask any questions. A total of 22 HV scans previously acquired are also available for use from ICHT if needed.

Participating HVs will undergo a single whole body MRI scan at RMH according to the trial specific scanning protocol. HV scans will be acquired in the following sequences (T1, fat/water, Dixon, ADC, etc) to mirror the clinical setting and on Siemens, Avanto and Aero (wide bore) MRI scanners (Supplementary S2 details sequences). Subjects with a larger Body Mass Index will be scanned on the Siemens Aero which has a larger bore diameter to optimise comfort.

Scan Acquisition– Myeloma Patients

Previously acquired patient scans will be identified by the investigators within the Sponsor's myeloma clinical service (between 2011 and 2020), supplemented by scans from ICHT, until the required sample size is reached. Scans will normally include the following sequences; T1, fat/water, Dixon, ADC, etc, and on the following MRI machines; Siemens, Avanto and Aero MRI scanners (Supplementary S2 for sequence details).

Scan Classification and Allocation to Study Phase

Patient scans will be categorised by the expert reference panel as showing inactive disease, active focal, active diffuse (focal or diffuse) and new disease. HV scans will be classified as normal (i.e. non-diseased). Scans will be allocated to Phase 1 or 2 as per Table 3. To minimise bias or 'over-learning', no more than 5 scans from the same patient will be allocated to Phase 1. Phase 2 scans will not include any patient scans that have been used in Phase 1 and thus comprise only those previously unseen by the ML algorithm. A subset of scans from phase 1 and 2 will be used to further train the algorithm at the start of Phase 3. Phase 3 validation scans have previously been acquired for the iTTiM trial (NCT02403102) and include a unique series of paired scans, previously unseen by the ML algorithm.

Scan Curation (Quality Control) and Anatomical Segmentation

Eligible scans will be curated immediately prior to transfer to an online platform for secure storage (ICR XNAT). This will ensure the ML algorithm is able to interpret all scans consistently. Curation scripts will be written in python and ensure that scans exhibit consistent characteristics such as:

correct sequential display of images, no missing slices, noting presence of unusual artifacts that might interrupt ML reads and other factors which might compromise interpretation. Further details on the data curation will be published elsewhere.

Phase 1 scans will then be manually segmented into 16 bone regions (Table 1) using a boundary box approach. These scans will be used to teach the ML algorithm to recognise active myeloma disease (focal or diffuse) and precision metrics will be evaluated in order to achieve the optimal algorithm. Initially, scans will be classified by the ML algorithm at scan level (i.e. patient level) only.

Testing of ML Algorithm – Radiology Reading Process

The ML algorithm will be tested by both experienced and inexperienced radiology readers.

Phase 2 scans will be subjected to the ML algorithm which will provide an ML overlay on all scans indicating areas of disease by means of a heat map. For each scan, a 'standard' and 'machine learning' version will be available. The trial statistician will randomly allocate reads to each of the (approximately 15 – 20) readers, using trial-specific algorithms written using Stata software (StataCorp, Texas). The reads will be performed in two batches to incorporate a wash-out period. Each batch will have 50% of cases with ML support and 50% without, to avoid reader training bias. The reading process will be described in a Reader Manual and all readers will receive appropriate training in viewing scans using the Biotronics 3D Web-based platform and completing a Read CRF available via Microsoft Forms (see supplementary file S3a). In the case of "inexperienced" readers, training will comprise a review of the CRFs and the viewing software with a basic training on reporting lexicon. A scribe will be provided to assist readers during the reading process and input data to the CRF in each batch of reads. Following a 4-week wash out period, readers will be presented with the second batch of reads with the opposite reading paradigm with regards to the ML support. The same cases will be allocated to the same readers. A subset of approximately 50 scans will be read a second time by a different reader as an interrater check.

In Phase 3, scans from the iTIMM study, comprising paired baseline and follow-up post treatment scans, will be used to test whether the ML algorithm is capable of distinguishing change in disease status (i.e. disease burden) between the two time-points. Reads will again be randomly allocated to the readers by the trial statistician. Readers will follow similar procedures to that outlined above with one set of paired scans having the ML overlay and the other with no ML overlay (for CRF see supplementary file S3b). A 4-week wash-out period will again apply between the two batches of reads. A subset of approximately 20 scans will be read a second time by a different reader as an interrater check.

Data collection

Reader responses will be captured using MS Forms with responses being transferred directly to an excel spreadsheet. Examples of the CRFs to be used in both ML validation phases are given as supplemental files (S3a, S3b). All readers will be provided with a manual describing CRF completion (including a lexicon of disease definitions) and use of the software viewing tools and overlay of the ML output heatmap and opportunity for live training using the online platform.

Outcome measures

Phase 1 – ML Algorithm Training Phase

Primary: Sensitivity for the detection of active myeloma on WB-MRI + ML detection tool against the reference standard.

Secondary: 1. Specificity; 2. F1 score (a single measure of precision and recall).

Phase 2 – ML Algorithm Clinical Testing Phase (Presence /Absence of active myeloma)

Primary: Difference in sensitivity of WB-MRI -/+ ML detection tool to diagnose the presence of active myeloma on a per-patient basis, by experienced readers, assessed against the reference standard.

Secondary: For comparison of WB-MRI -/+ML: 1. Per-site sensitivity to diagnose active disease; 2. Reading time; 3. Specificity; 4. Agreement with reference standard to categorise disease as focal, diffuse and/or extramedullary; 5. Sensitivity of non-experienced readers for presence of active disease.

Phase 3 – ML Algorithm for quantification of disease burden with clinical testing

Primary : agreement between experienced readers and the reference standard in scoring overall burden of disease before and after treatment for response categorisation -/+ ML quantification tool.

Secondary: For comparison of WB-MRI -/+ML: 1. Reading time; 2. Agreement of categorisation of patients as responder or non-responder with the reference standard; 3. Agreement of non-experienced readers for burden of disease and categorisation of response; 4. Estimated difference in cost for radiology reading time for WB-MRI -/+ML.

Proposed tertiary: Verification of the team's previously published work regarding reverse classification accuracy: predicting segmentation performance in the absence of a reference standard (21).

Sample size

Phase 1:

We will train the ML algorithm on a set of scans without and with active disease that will reflect the categories of disease that may be encountered in clinical practice. The number of cases used for training is arbitrarily chosen reflecting the knowledge that a large number of training datasets will improve training accuracy, counterbalanced with the resources needed to curate and annotate a large number of datasets.

Phase 2:

The study is powered on the primary outcome of sensitivity.

In a meta-analysis, Wu et al have reported a pooled sensitivity of 88% and a pooled specificity of 86% (0.86 for WB-MRI with DW-MRI) (8). We anticipate that the addition of ML could increase this by at least 7.5%, from 88% to 95.5%. There is no background data to indicate the expected proportion of discordant pairs so we have estimated this as (1-0.955)*0.88 + 0.955*(1-0.88), which is equal to 0.154. To achieve 80% power using a two-sided alpha of 0.05 would require a total of 203 patients positive for myeloma using the gold standard.

If it is assumed that the specificity will be unchanged using ML, a total number of cases with no active disease of 150 (50 HV, 100 inactive treated myeloma), will give 80% power to show that the difference is above a non-inferiority limit of 10%.

Phase 3 training:

Approximately 200 cases that have at least two time points will be taken from phase 1 and 2, with active disease present at least at one time point, and used for training and validation for burden of disease; this will ensure efficient use of all data and segmentations.

Phase 3 clinical testing:

This sample size is fixed at 60 patients, the full sample size of the iTIMM study, each of whom has a baseline and one post treatment scan.

Statistical Analysis

Phase 1 analysis

The ability to correctly localise and detect active disease will be evaluated by calculating sensitivity, specificity and the F1 score (a single measure of precision (positive predictive value) and recall (sensitivity)) for multiple algorithms and compared against the reference standard. Following Trial Steering Committee (TSC) approval, the optimal algorithm will move forward to phase 2.

Phase 2 analysis

In phase 2, the percentage of patients with active disease on WB-MRI +/- ML support who have positive reference standard will be compared using McNemar's test with a two-sided alpha of 0.05. Per patient and per site sensitivity and specificity with and without ML support will be reported with 95% confidence intervals. Reading time will be compared using Wilcoxon's test for paired data and described using summary statistics.

The same analysis of sensitivity, specificity and reading time will be repeated for inexperienced readers.

Agreement between experienced and inexperienced readers will be measured in a subset of cases with a Kappa coefficient, and overall proportion of concordant cases.

All other endpoints will be summarised using descriptive statistics.

Although the study is powered to detect superiority of the primary endpoint, if sensitivity is shown to be non-inferior using ML and reading time is both clinically and statistically significantly lower using ML, this would be considered as an indication to proceed. Non-inferiority in this context will be defined as having any possible reduction in sensitivity with ML significantly higher than a lower limit of -10% (using Tangos' test with one-sided alpha 0.05)

Phase 3 analysis

In phase 3, the difference between the experienced readers' disease score to the reference standard disease score will be recorded and compared +/- ML support using Wilcoxon's test. Differences from scores given by experienced readers and the reference standard will be described using Bland-Altman plots for scores +/- ML support.

All other endpoints will be summarised using descriptive statistics.

A simple cost-effectiveness analysis may be performed depending on study findings, such as the reading time.

Procedure(s) to account for missing or spurious data

If a scan is incomplete or the file is corrupted and not evaluable, it will be excluded from the dataset. If a set of radiology reads is incomplete, a new trained reader will be identified to do the full allocation of reads.

Timing and responsibility for analyses

Analyses will take place at both the end of phase 2 and then again at the end of phase 3, when all readings have been completed.

Patient and public involvement (PPI)

A PPI representative was appointed from an established group at Myeloma UK. The individual gave in-depth feedback on the study, particularly on the relevance to patient care and the use of retrospective patient data and HV scans. Myeloma UK is fully supportive of the project and is willing to assist with dissemination of important findings to the Myeloma UK community.

Safety

As this study is recruiting healthy volunteers only, an a-priori agreement has been reached with the Sponsor that safety reporting is not required. Sponsor procedures in respect of incidental (i.e. unexpected) findings in healthy volunteers will be adhered to and results captured within the Trust's Clinical Record.

Monitoring against Source Data will not be required which is in line with the Sponsor's policy on non-CTIMP (Clinical Trial of Investigational Medicinal Product) trials.

Trial funding, organisation and administration

The study has been awarded funding by Medical Research Council (MRC) NIHR EME Awards Body (NIHR EME Project ID: 16/68/34). In addition, the Department of Radiology has agreed to fund the cost of HV WB-MRI scans. The cost of recruitment and consenting of HVs will be requested through

the NHS CRN. RMH is the study sponsor responsible for initiating and managing the study and the coordinating centre, including sign-off of the study protocol.

A Trial Management Group (TMG) meeting will be held regularly to ensure satisfactory progress of the study. A TSC will provide independent oversight for the study, review the development of the ML algorithm, and advise the TMG where problems may arise. The TSC will include a Patient Advocate.

Ethics and dissemination

Ethical approval for MALIMAR was granted on 21/11/2017 (REC) and 21/12/2017 (Health Research Authority) Here, we report version 3.0 of the protocol. All participating sites gained local approval prior to study participation.

Any protocol modifications will be submitted for approval to the REC, reflected in the online registration and disseminated by e-mail to site principal investigators and trial coordinators. The statistician will have access to the final linked trial dataset. There are no plans to provide public access to the full protocol, participant-level data, or statistical code. The researchers aim to publish results in a peer-reviewed journal and share via social media and conferences. Authorship will be determined according to academic standards.

Discussion

This study aims to develop and validate a ML algorithm to augment the performance and efficiency of the radiology reading process using WB-MRI. The results will show the impact of using the ML tool and outcomes of the study will have implications for the application of ML with WB-MRI in myeloma patients across the NHS. It is anticipated that feasibility analysis will follow the successful completion of this study to pilot the implementation of the ML tool in a real-time prospective study prior to future clinical setting.

To avoid bias we ensure: 1) comparator and intervention tests are read by readers that are fully blinded to the reference standard; 2) a mixture of cases with and without disease; 3) the reads will be presented such that radiologists must read a mixture of cases without or with ML support during each round of reading including a wash-out period. We will have unavoidable incorporation bias, as the expert reference panel will use the MRI as part of the reference standard. The reference panel will consist of a single person's opinion which is a limitation to our study. If resources had allowed, the gold standard would have been to have two blinded opinions with a consensus panel in cases of disagreement. Other limitations include varying scan quality as data is acquired over a 9-year period; and replicating clinical reporting in a retrospective study setting can be challenging.

In conducting this study, we will have acquired possibly the largest set of characterised myeloma patient MRI scans in the UK and we anticipate this will form the basis of a unique training resource in the future.

Machine learning techniques in WB-MRI scans of patients with myeloma is likely to be transferable to other malignancies. In prostate and breast cancer, quantification of metastatic bone disease is an unmet need as bone only disease is not uncommon and is currently classified as non-measurable by RECIST 1.1 (22). The participating HVs will be consented to allow the anonymised datasets to be a future resource for the wider research community.

Study status

The MALIMAR study opened on 26 April 2018 using protocol version 1.0 (30 Oct 2017). The study was in phase II, using protocol version 3.0 (31 Jan 2019), at date of submission. Protocol amendments are documented in Supplementary S4.

Author contributions:

AR, CM, TaB, BG, SW, TQ, ThB, SD, ML, MK and DK conceptualization and methodology; AR, CM, TaB, ThB, MK, BG, TQ, XF and SW investigation; EA and AR resources; ThB and SD data curation; LS and EG formal analysis; AR, DK and CM supervision; LS writing – original draft; AR,

CM, TaB, BG, LW and LS writing – review and editing; BG, TQ, XF and SW data visualisation; LW project administration; EA and AR funding acquisition.

Competing interests:

AR receives honoraria for educational lecture at Garmisch International Symposium, has an unpaid role on the European Society of Radiology Board of Directors and receives travel cost support where necessary. BG receives grants from other entities; EU commission and UKRI London Medical Imaging & Artificial Intelligence Centre for Value Based Healthcare, is a Scientific advisor for Kheiron Medical Technologies (Jan 2018 – Sep 2021) and receives stock options as part of standard employment packages from both Kheiron Medical Technologies and HeartFlow. EA has a patent pending for Machine Learning in Alzheimer's disease and has a role on the scientific advisory board for Radiopharm Theranostics Limited. MK receives grants from both Myeloma UK and Celgene/BMS, and consulting fees or payments from AbbVie, BMS/Celgene, Janssen, GSK, Karyopharm, Takeda and Seagen. CM & DK receive additional funding as a co-investigator on a radiology NIHR study and is part of the joint venture Celescan with the Royal Marsden, The Institute of Cancer Research and Sopra Steria. TB receives additional funding from CRUK grant funding (NCITA) and NIHR (HTA) and receives honoraria from Bayer.

Funding:

This study (ID: 16/68/34) is funded by the Efficacy and Mechanism Evaluation (EME) Programme, an MRC and NIHR partnership. In addition, the Department of Radiology has agreed to fund the cost of healthy volunteer whole body MRI scans. The cost of recruitment and consenting of healthy volunteers will be requested through the NHS Clinical Research Network. The views expressed in this publication are those of the authors and not necessarily those of the MRC, NHS, the NIHR, or the Department of Health and Social Care.

EG and LS's posts are part funded by the National Institute for Health and Care Research (NIHR) Biomedical Research Centre at The Royal Marsden NHS Foundation Trust and the Institute of Cancer

Research, London. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

SW is supported by the UKRI London Medical Imaging & Artificial Intelligence Centre for Value Based Healthcare.

Data sharing statement:

Anonymised data are available upon reasonable request as a resource for the wider research community from Andrea Rockall (ORCID ID: 0000-0001-8270-5597), providing consent has been granted from all participants.

References

1. NICE. Myeloma: diagnosis and management NICE guideline [NG35]. 2016.
2. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. The lancet oncology. 2014;15(12):e538–48.
3. Dimopoulos MA, Hillengass J, Usmani S, Zamagni E, Lentzsch S, Davies FE, et al. Role of magnetic resonance imaging in the management of patients with multiple myeloma: a consensus statement. 2015;
4. Pearce T, Philip S, Brown J, Koh DM, Burn PR. Bone metastases from prostate, breast and multiple myeloma: differences in lesion conspicuity at short-tau inversion recovery and diffusion-weighted MRI. Br J Radiol. 2012;85(1016):1102–6.
5. Squillaci E, Manenti G, di Stefano F, Miano R, Strigari L, Simonetti G. Diffusion-weighted MR imaging in the evaluation of renal tumours. Journal of Experimental and Clinical Cancer Research. 2004;23(1):39–46.
6. Dutoit JC, Vanderkerken MA, Anthonissen J, Dochy F, Verstraete KL. The diagnostic value of SE MRI and DWI of the spine in patients with monoclonal gammopathy of undetermined significance, smouldering myeloma and multiple myeloma. Eur Radiol. 2014;24(11):2754–65.
7. Messiou C, Hillengass J, Delorme S, Lecouvet FE, Moulopoulos LA, Collins DJ, et al. Guidelines for acquisition, interpretation, and reporting of whole-body MRI in myeloma: myeloma response assessment and diagnosis system (MY-RADS). Radiology. 2019;291(1):5–13.
8. Wu L, Gu H, Zheng J, Xu X, Lin L, Deng X, et al. Diagnostic value of whole-body magnetic resonance imaging for bone metastases: a systematic review and meta-analysis. Journal of Magnetic Resonance Imaging. 2011;34(1):128–35.
9. Messiou C, Porta N, Sharma B, Levine D, Koh DM, Boyd K, et al. Prospective evaluation of whole-body MRI versus FDG PET/CT for lesion detection in participants with myeloma. Radiology: Imaging Cancer. 2021;3(5).

10. Pawlyn C, Fowkes L, Otero S, Jones JR, Boyd KD, Davies FE, et al. Whole-body diffusion-weighted MRI: a new gold standard for assessing disease burden in patients with multiple myeloma? Leukemia. 2016;30(6):1446–8.

11. Mateos MV, Hernández MT, Giraldo P, de la Rubia J, de Arriba F, Corral LL, et al. Lenalidomide plus Dexamethasone for High-Risk Smoldering Multiple Myeloma. New England Journal of Medicine. 2013 Aug;369(5):438–47.

12. Padhani AR, Koh DM, Collins DJ. Whole-body diffusion-weighted MR imaging in cancer: current status and research directions. Radiology. 2011;261(3):700–18.

13. Radiologists RC of. Clinical radiology UK workforce census 2015 report. The Royal College of Radiologists London; 2016.

14. Juntu J, Sijbers J, de Backer S, Rajan J, van Dyck D. Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine. 2010;31(3):680–9.

15. Pauly O, Glocker B, Criminisi A, Mateus D, Möller AM, Nekolla S, et al. Fast multiple organ detection and localization in whole-body MR Dixon sequences. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2011. p. 239–47.

16. Lavdas I, Rockall AG, Castelli F, Sandhu RS, Papadaki A, Honeyfield L, et al. Apparent diffusion coefficient of normal abdominal organs and bone marrow from whole-body DWI at 1.5 T: the effect of sex and age. American Journal of Roentgenology. 2015;205(2):242–50.

17. Lavdas I, Glocker B, Rueckert D, Taylor SA, Aboagye EO, Rockall AG. Machine learning in whole-body MRI: experiences and challenges from an applied study using multicentre data. Clin Radiol. 2019;74(5):346–56.

18. Lavdas I, Glocker B, Kamnitsas K, Rueckert D, Mair H, Sandhu A, et al. Fully automatic, multiorgan segmentation in normal whole body magnetic resonance imaging (MRI), using classification forests (CF s), convolutional neural networks (CNN s), and a multi-atlas (MA) approach. Med Phys. 2017;44(10):5210–20.

19. Giles SL, Desouza NM, Collins DJ, Morgan VA, West S, Davies FE, et al. Assessing myeloma bone disease with whole-body diffusion-weighted imaging: comparison with x-ray skeletal survey by region and relationship with laboratory estimates of disease burden. Clin Radiol. 2015;70(6):614–21.

20. Giles SL, Messiou C, Collins DJ, Morgan VA, Simpkin CJ, West S, et al. Whole-body diffusion-weighted MR imaging for assessment of treatment response in myeloma. Radiology. 2014;271(3):785–94.

21. Valindria V v, Lavdas I, Bai W, Kamnitsas K, Aboagye EO, Rockall AG, et al. Reverse classification accuracy: predicting segmentation performance in the absence of ground truth. IEEE Trans Med Imaging. 2017;36(8):1597–606.

22. Lecouvet FE, Talbot JN, Messiou C, Bourguet P, Liu Y, de Souza NM, et al. Monitoring the response of bone metastases to treatment with Magnetic Resonance Imaging and nuclear medicine techniques: a review and position statement by the European Organisation for Research and Treatment of Cancer imaging group. Eur J Cancer. 2014;50(15):2519–31.

Figure 1: MALIMAR Study Flow Diagram

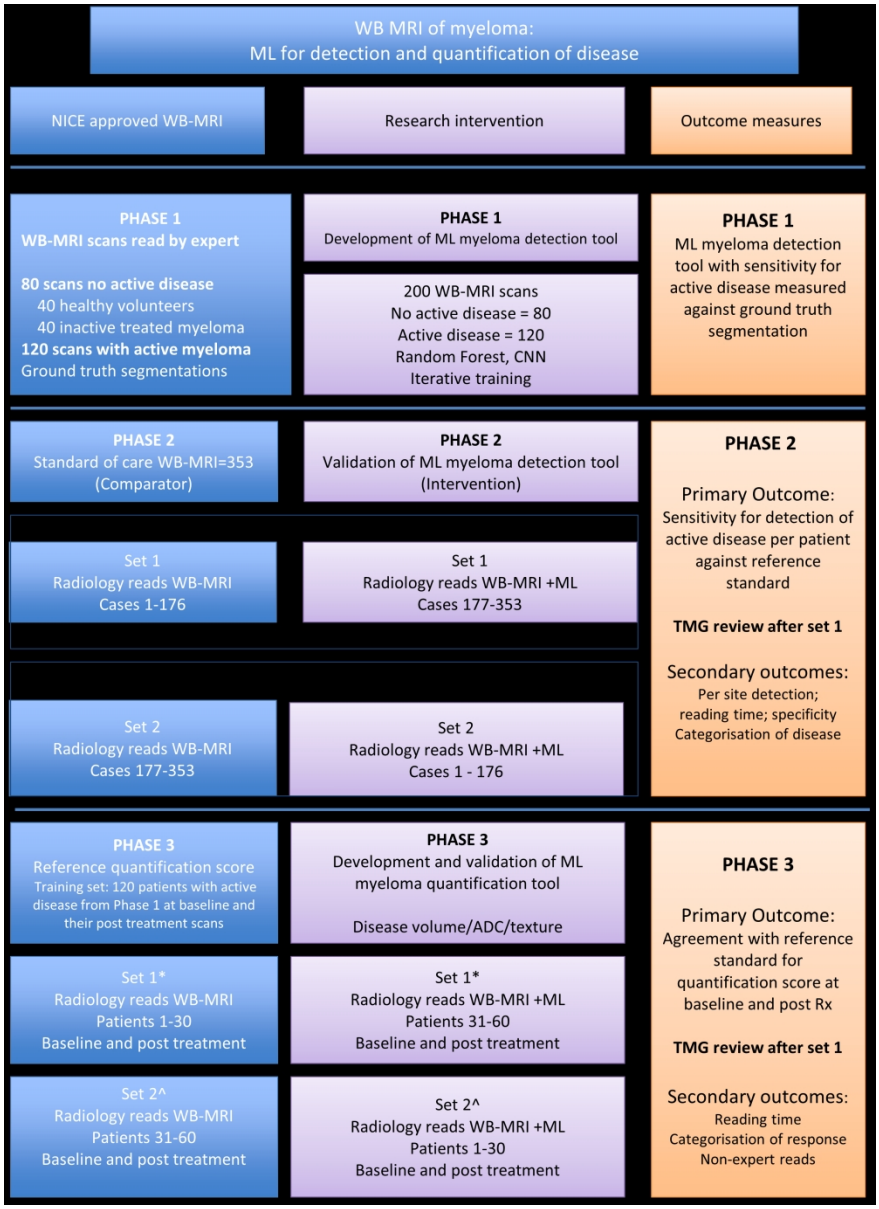

Figure 1: Study Flow Chart

526x723mm (130 x 130 DPI)

The ROYAL MARSDEN
NHS Foundation Trust

The MALIMAR Study Healthy Volunteer Consent Form

Study Reference Numbers: CCR 4820: IRAS No.: 233501

NHS No. ☐☐☐☐☐☐☐☐☐☐

Healthy volunteer Trial ID: ☐☐☐☐☐☐☐☐☐

Name of Lead Researcher: ……

Please initial box

1. I confirm that I have read and understand the Healthy Volunteer Information Sheet version 2.0 dated 07/12/18 for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐
3. If I request withdrawal from the study, I give permission that my data already collected within the study can be anonymised and used. ☐
4. I understand that relevant sections of my medical notes may be looked at by responsible individuals from the research team, from regulatory authorities or from the NHS trust where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
5. I consent to undergo an MRI scan under the supervision of the responsible clinician for this research. I understand that if any health related issues come to light as a result of undergoing this scan, otherwise known as 'incidental findings', that I and my General Practitioner will be promptly informed of these issues. ☐
6. I agree to participate in the MALIMAR study. ☐
7. I give permission for the data collected during the study to be used in further ethically approved research within and outside the UK in the field of imaging research. I understand this will not include any personal data from which I could be identified. ☐

Name of Healthy Volunteer	Date	Signature
Name of person taking consent (PI or approved signatory)	Date	Signature

Original for Investigator's Site File; 1 copy for volunteer; 1 copy for hospital notes; 1 copy to be sent to RM-CTU

The ROYAL MARSDEN
NHS Foundation Trust

MALIMAR

Healthy Volunteer Information Sheet

Development of machine learning support for reading whole body diffusion weighted magnetic resonance imaging (WB-DW-MRI) in myeloma for the detection and quantification of the extent of disease before and after treatment.

Short Title: MAchine **L**earning **I**n **M**yelomA **R**esponse

7th December 2017

Version 2.0

CCR Number: 4820

IRAS (Integrated Research Application System) No. 233501

You are being invited to take part in a research study. Before you decide whether or not to take part it is important for you to understand why we are doing this research and what it involves. Please take time to read the following information carefully and discuss it with relatives, friends, and your GP if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time deciding whether or not you wish to take part.

You can learn more about clinical research on the Cancer Research UK's patient website (www.cancerhelp.org.uk)

Invitation

If you are 40 years or above the Radiology Department at the Royal Marsden hospital would like to invite you to take part in a research study. This will involve you having a particular type of Magnetic Resonance Imaging (MRI) scan known as a Whole-Body Diffusion Weighted MRI scan or 'WB-DW-MRI'.

Before you decide to participate it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with friends, relatives and your GP if you wish. Ask us if there is anything that is not clear or if you would like more information. Thank you for reading this Information Sheet.

What is the purpose of the study?

There have been enormous advances in recent years in the technology used to take pictures (images) of the internal anatomy of cancer patients to better identify sites of disease. These images (or scans) can now provide a more accurate indication of the scope or spread of disease. They can also be used for assessing disease response to different drugs or treatments.

MRI (magnetic resonance imaging) has the advantage over other types of scanning (e.g. computerised tomography or 'CT') in that it does not involve the delivery of any radiation dose. In particular, a new type of MRI, called Whole Body Diffusion Weighted MRI (WB-DW-MRI) can provide especially precise images of diseased compared to healthy tissues. As a result, it is now being more widely used in cancer treatment centers throughout the world.

Despite these advantages, WB-DW-MRI has an important disadvantage. Each scan is made up from over a thousand images, each of which needs to be read and interpreted by an expert Radiologist. Thus, the time taken to read a single WB-DW-MRI scan is much longer than for a normal MRI scan, meaning that few NHS treatment centres (or hospitals) are able to offer them to patients.

Members of the research team from the Royal Marsden Hospital and Imperial College London have already undertaken some work to ascertain how computers can reduce the time taken to read WB-DW-MRI scans. The technique is called 'machine learning' and basically teaches a computer to detect areas of suspicion or concern for disease on WB-DW-MRI scans. The 'trained' computer can then make an initial and very rapid interpretation of the images taken during a scan. These images can then be presented to the expert radiologist to make the final interpretation. In addition to training computers to read scans more quickly, we also want to train computers to interpret differences between scans taken from the same patient at different time-points. This will allow us to accurately assess change in disease extent or response to treatment over time.

However, in order to train the computers, we need examples of WB-DW-MRI images taken from both diseased (cancerous) and healthy tissues. In this study we are concentrating on patients with myeloma (cancer of white blood cells). We have already acquired WB-DW-MRI images from many patients with this type of cancer. **So now, we are seeking your help to acquire WB-DW-MRI images from healthy tissues for the M**achine **L**earning **I**n **M**yelom**A R**esponse (**MALIMAR**) study**.**

What will happen to me if I decide that I would like to take part?

Before we can enter you to the study, we will need to check that you can have an MRI scan and that you are suitable to take part. Some people cannot have an MRI scan. These include people with a pacemaker, metal heart valves, aneurysm clips in the brain or people who have had metal fragments in their eyes. In addition, we are unable to include volunteers who have had or have a significant illness as this may affect the scan.

It may also not be appropriate for you to take part if you have had extensive surgery previously. Our study researcher will confirm these points with you before you are admitted to the trial. As advised above we are only recruiting volunteers aged 40 and above: anyone under this age will have to be

excluded from participating because they will not be a suitable comparator. Once we have confirmed that you are suitable to enter the trial, we will ask you to sign an Informed Consent Form and then book your scan. Some volunteers may be asked to attend early evening or week-end appointments to avoid busy times during the day when the MRI Unit is reserved for patients. There are usually no special preparations and no injection or drugs will be given. All instructions for the scan will be in your MRI appointment letter. When you come for the scan you are advised to wear clothing without metal fastenings and to avoid using make-up or mascara. You can wear glasses, but will need to take these off during the scan. A locker will be provided for your valuables.

The MRI scan will be carried out by radiographers who are trained to carry out the scans. MRI uses a magnetic field and radio waves to build up detailed images of your internal anatomy by detecting signals sent out by water molecules. It is not painful, but you will have to lie still for the duration of the scan which can be up to 60 minutes. The scanner produces a variety of loud noises during the scan which are made by the magnetic coils that switch on and off during the scan. These are important in measuring the signals from your body to create the images. They are switched on and off very quickly and they vibrate, which is what causes the noise.

Some people may find the noise level uncomfortable and the table quite hard to lie on. You will be provided with earplugs to help reduce the noise. The scanner is open at both ends, but some people may find it claustrophobic. During the scan the radiographer can see you from the control room and can talk to you through an intercom. You will be given a call button to press to alert attention and can listen to music during the scan. You can leave as soon as your scan is finished and can eat and drink as normal. There are no side effects from the MRI scan itself.

Why am I being invited to take part?

You will be reading this Information Sheet because you have responded to one of our advertisements for Healthy Volunteers to take part. If we invite you to sign a Consent Form then you are eligible to take part in the study. If you are not eligible to participate we will explain the reason.

Do I have to take part?

No, it is up to you to decide whether or not to take part. If you do choose to take part you will be asked to sign a consent form, a copy of which will be given to you for your records along with this information sheet about the study. Your legal rights are not affected by participation in the study.

What happens if I change my mind during the study?

Your participation in this study is entirely voluntary. If you agree to take part and then change your mind and wish to withdraw, you may do so at any time. If you decide to not join the study or to discontinue in the study, this will not affect any future care or treatment you receive.

What are the risks and the benefits of taking part in this study?

A possible risk in taking part is a degree of discomfort you may encounter in undergoing the MRI scan. As we said above, unlike other forms of imaging (e.g. CT scans) MRI does not deliver radiation and no drugs or other medication will be given. You will be registered on the Royal Marsden Hospital Information System and a report of your scan results will be held on this system. If an unexpected finding of concern is discovered, a doctor will call you to discuss your scan report. We will also send a copy of the report to your GP who will then advise you regarding any follow-up investigations that may be needed. This could lead to some anxiety. If unexpected findings are discovered which are not concerning, we will send you a letter to explain the findings and copy this letter to your GP. You may then wish to call us or your GP for more information. If there are no unexpected findings we will not contact you or your GP.

In general, the research will not be of direct benefit to you, but may prove to be of benefit to others in the future. However, possible benefits are that you may find it satisfying to have contributed to medical research and, should an unexpected finding be discovered you may feel that the early detection and diagnosis will result in a better outcome. If you wish to have a copy of your scan report, you may ask for this.

What if something goes wrong?

It is unlikely that anything will go wrong but, if you wish to complain, you can do so using the normal NHS complaints procedure. If taking part harms you in any way, there are no special compensation arrangements, but the hospital would be liable for any negligence on the part of hospital staff. Your legal rights are not affected by giving your consent to participate in this study.

Who is organizing and funding the research?

This study is being organised by The Royal Marsden NHS Foundation Trust with participation from The Institute of Cancer Research, Imperial College London and Imperial Healthcare NHS Foundation Trust. The study is being funded by a National Institute for Health Research grant as part of their Efficacy and Mechanism Evaluation programme.

Will my taking part in this study be kept confidential?

1) Clinical Information: You will need to be given a Royal Marsden hospital number in order to receive the WB-DW-MRI scan. The resulting scan report will be held on our clinical Hospital Information (NHS PACS) System which is the system we use for holding all NHS patient information. Access to this system is subject to the normal Trust-based information governance controls. If, in the event of unexpected findings, you require further diagnostic investigations, your GP will be informed and your scans and accompanying data will be made available to the hospital treating you.

2) Research Information: Your scan data will be anonymised and identified by a unique trial identification number. Your unique trial number will be used to make sure you cannot be identified by members of the research team that are not part of the NHS staff at RMH. The data from your scan which will be used in the MALIMAR study will only be available to authorised members of our research team so they can collect information needed for this research study and also to check that it is correct. All information will be kept confidential, and your name, date of birth and other identifiable information will be removed from your scans prior to archiving. We will also ask you to consent to allow your data that has been collected in the study to be sent outside of the UK and to be used in future ethically approved studies. This information will not include any personal information that could directly identify you.

What will happen to the results of this study?

As soon as there are reliable results, they will be published in a respected peer reviewed medical journal and presented in various scientific meetings. Your identity will not be revealed in any report, publication or presentation. The results will be available on request.

How is the trial monitored for safety?

This study has been carefully planned by leading cancer specialists and approved by the Oxford C Research Ethics Committee (REC), the Royal Marsden Hospital Committee for Clinical Research (CCR) and the Health Research Authority (HRA). The members of the study team will be meeting at regular intervals to monitor the progress and safety of the study. Full (100%) monitoring will be carried out to ensure that where incidental findings come to light, both you and your GP are promptly informed.

What do I do now?

We would be happy to answer any questions you may have about the study. You can telephone us, or speak to us again. Please discuss this information with your family, friends or your GP if you wish. If you require further information about this study please contact:

Professor Andrea Rockall,
Chief Investigator,
Clinical Chair Radiology,
ICTEM Building,
Imperial College Healthcare HNS Trust,
Du Cane Road
London, W12 0NN
Tel: 0207 59 42792 (Personal Assistant to Professor Rockall)

Dr Christina Messiou,
Principal Investigator,
Consultant Radiologist,
The Royal Marsden NHS Foundation Trust,
Fulham Road
London, SW3 6JJ
Tel: 0208 661 3216

Veronica Morgan
MRI Research Superintendent Radiographer
Clinical Magnetic Resonance Unit, Sutton
The Royal Marsden NHS Foundation Trust
Tel 02089156493

Thank you for reading and considering taking part in this study.

Funding Acknowledgement: Funding from the National Institute for Health Research – Efficacy and Mechanism Evaluation (NIHR – EME) programme for the MALIMAR study is acknowledged.

Participant Type	Study name	Site	MRI Machine Name	Sequences acquired
Healthy Volunteers	MALIMAR	Royal Marsden	Siemens Aera	Haste localiser, Sag T1, T2 whole spine, DWI (B50,600,900 and ADC), fl3d_CAIPI_wb_tra_BH_20 and T2 HASTE Vertex to knees
Healthy Volunteers	MALIMAR	Royal Marsden	Siemens Avanto	localiser, Sag T1, T2 whole spine, DWI (B50,600,900 and ADC), fl3d_vibe_dixon_TRA_15deg 256_pocS and T2 HASTE Vertex to knees
Myeloma Patients	MALIMAR	Royal Marsden	Siemens Aera	Haste localiser, Sag T1, T2 whole spine, DWI (B50,600,900 and ADC), fl3d_CAIPI_wb_tra_BH_20 and Vertex to knees
Myeloma Patients	MALIMAR	Royal Marsden	Siemens Avanto	localiser, Sag T1, T2 whole spine, DWI (B50,600,900 and fl3d_vibe_dixon_TRA_15deg 256 Vertex to knees
Myeloma Patients	MALIMAR	ICHT	Siemens Aera	Axial dixons x 4 350 slices each (total: 1400) B 50 248 slices B900 248 slices ADC 248 slices Sag T1 spine 15 slices Sag T2 spine 15 slices

MALIMAR Radiology Reads - CRF Phase 2

Version 4, 06 September 2021

* Required

1. Scan ID *

2. Reader ID *

3. Round *

- ◯ Round 1
- ◯ Round 2

4. Date of Read *

Please input date (dd/MM/yyyy)

5. Start time of read - Enter in format: HH:MM using 24 hour clock *

6. Disease status - BONES - Record Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Cervical Spine	◯	◯	◯	◯
Dorsal Spine	◯	◯	◯	◯
Lumbar Spine	◯	◯	◯	◯
Pelvis	◯	◯	◯	◯
Long Bones	◯	◯	◯	◯
Skull	◯	◯	◯	◯
Ribs / Clavicles / Sternum / Scapulae	◯	◯	◯	◯

7. Disease status - BONES - Record maximum size of Active / Focal lesions (mm) *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Cervical Spine	○	○	○	○
Dorsal spine	○	○	○	○
Lumbar spine	○	○	○	○
Pelvis	○	○	○	○
Long Bones (max. long axis)	○	○	○	○
Skull	○	○	○	○
Ribs / Clavicles / Sternum / Scapulae (max. long axis)	○	○	○	○

8. Disease Status - BONES - How confident are you in your assessment of Active / Focal lesions *

	Not at all confident	Some confidence	Confident	Very Confident
Cervical Spine	○	○	○	○
Dorsal Spine	○	○	○	○
Lumbar Spine	○	○	○	○
Pelvis	○	○	○	○
Long Bones	○	○	○	○
Skull	○	○	○	○
Ribs / Clavicles / Sternum / Scapulae	○	○	○	○

9. Disease Status - Record if diffuse disease was present at any of these sites? *

	Yes	No
Cervical Spine	◯	◯
Dorsal Spine	◯	◯
Lumbar Spine	◯	◯
Pelvis	◯	◯
Long Bones	◯	◯
Skull	◯	◯
Ribs / Clavicles / Sternum / Scapulae	◯	◯

10. How confident were you in your assessment of diffuse disease at these sites? *

	Not at all confident	Some confidence	Confident	Very confident
Cervical Spine	◯	◯	◯	◯
Dorsal Spine	◯	◯	◯	◯
Lumbar Spine	◯	◯	◯	◯
Pelvis	◯	◯	◯	◯
Long Bones	◯	◯	◯	◯
Skull	◯	◯	◯	◯
Ribs / Clavicles / Sternum / Scapulae	◯	◯	◯	◯

11. Was extramedullary disease present at any site? *

◯ Yes

◯ No

12. If extramedullary disease was present at any site - state location(s) separated by a semi-colon

13. If extramedullary disease was present, what was your level of confidence in assessing this? *

- ◯ Not confident at all
- ◯ Some confidence
- ◯ Confident
- ◯ Very confident
- ◯ Not Applicable, no extramedullary disease is seen.

14. Confidence in assessing overall disease status on this scan (i.e. in determining the presence or absence of ANY active disease) *

- ◯ Not confident at all
- ◯ Some confidence
- ◯ Confident
- ◯ Very confident

15. Stop time of read - RECORD IMMEDIATELY AFTER COMPLETING CLINICAL READ - Enter in format: HH:MM using 24 hour clock *

16. TO BE COMPLETED FOLLOWING THE CLINICAL READ:
Was a Machine Learning Image available *

- ◯ Yes
- ◯ No

17. If a Machine Learning 'ML' Image was available, please indicate whether sites were positive for active / focal disease, i.e. was there an ML finding?

	Highly likely negative on ML	Probably negative on ML	Probably positive on ML	Highly likely positive on ML
Cervical Spine	◯	◯	◯	◯
Dorsal Spine	◯	◯	◯	◯
Lumbar Spine	◯	◯	◯	◯
Pelvis	◯	◯	◯	◯
Long Bones	◯	◯	◯	◯
Skull	◯	◯	◯	◯
Ribs / Clavicles / Sternum / Scapulae	◯	◯	◯	◯

18. If a Machine Learning 'ML' image was available, please indicate whether sites were positive for diffuse disease, i.e. was there an ML finding?

	Highly likely negative on ML	Probably negative on ML	Probably positve on ML	Highly likely positive on ML
Cervical Spine	○	○	○	○
Dorsal Spine	○	○	○	○
Lumbar Spine	○	○	○	○
Pelvis	○	○	○	○
Long Bones	○	○	○	○
Skull	○	○	○	○
Ribs / Clavicles / Sternum / Scapulae	○	○	○	○

19. Scan Quality: What was the quality of the WB-MRI used for this read? *

	Good	Adequate	Poor
1. B 900	○	○	○
2. ADC	○	○	○
3. T1 sequences	○	○	○

20. Please enter any specific comments you have on scan quality

21. Reader confirmation: My responses have been accurately reported on this CRF (enter 'yes' if in agreement with this statement) *

- ◯ Yes
- ◯ No

MALIMAR Radiology Reads - CRF Phase 3

Version 2, 31 March 2022

* Required

1. Scan ID Post Treatment Scan (PT) *

2. Scan ID - Baseline Scan (BL) *

3. Reader ID *

4. Phase 3 - Round *

- ◯ Round 1
- ◯ Round 2

5. Date of Read *

Format: M/d/yyyy

6. Start time of read - Enter in format: HH:MM using 24 hour clock *

7. CERVICAL SPINE - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

8. CERVICAL SPINE - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

9. CERVICAL SPINE - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

10. DORSAL SPINE - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

11. DORSAL SPINE - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

12. DORSAL SPINE - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

13. LUMBAR SPINE - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

14. LUMBAR SPINE - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

15. LUMBAR SPINE - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

16. PELVIS - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

17. PELVIS - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

18. PELVIS - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

19. LONG BONES - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	○	○	○	○
Baseline	○	○	○	○

20. LONG BONES - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

21. LONG BONES - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

22. SKULL - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	◯	◯	◯	◯
Baseline	◯	◯	◯	◯

23. SKULL - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	◯	◯	◯	◯
Baseline	◯	◯	◯	◯

24. SKULL - Was Diffuse Disease present? *

	Yes	No
Post Treatment	◯	◯
Baseline	◯	◯

25. RIBS / CLAVICLES / STERNUM / SCAPULAE - Number of Active / Focal Lesions *

	0	1 - 4	5 - 10	>10
Post Treatment	◯	◯	◯	◯
Baseline	◯	◯	◯	◯

26. RIBS / CLAVICLES / STERNUM / SCAPULAE - Maximum size (mm) of Active / Focal Lesions *

	<10mm	10 - 20mm	>20mm	Not Applicable, No Focal lesions seen at this site
Post-Treatment	○	○	○	○
Baseline	○	○	○	○

27. RIBS / CLAVICLES / STERNUM / SCAPULAE - Was Diffuse Disease present? *

	Yes	No
Post Treatment	○	○
Baseline	○	○

28. Was extramedullary disease present at any site? *

○ Yes

○ No

29. If extramedullary disease was present at any site - state location(s) separated by a semi-colon

30. If extramedullary disease was present, what was your level of confidence in assessing this? *

- ◯ Not confident at all
- ◯ Some confidence
- ◯ Confident
- ◯ Very confident
- ◯ Not Applicable, no extramedullary disease is seen.

31. OVERALL RESPONSE - Change in Disease Status (Baseline - Post-Treatment) *

	Complete Response	Partial Response	Stable Disease	Disease Progression
Response category	◯	◯	◯	◯

32. OVERALL RESPONSE - CONFIDENCE - How confident were you in assessing overall response? *

	Not at all confident	Some confidence	Confident	Very confident
Confidence category	◯	◯	◯	◯

33. Stop time of read - RECORD IMMEDIATELY AFTER COMPLETING CLINICAL READ - Enter in format: HH:MM using 24 hour clock *

34. TO BE COMPLETED FOLLOWING THE CLINICAL READ:
Was a Machine Learning Image available *

- ◯ Yes
- ◯ No

35. If Machine Learning 'ML' Images were available, please indicate category of response suggested by ML

	Complete Response	Partial Response	Stable Disease	Progressive Disease
Response category	◯	◯	◯	◯

36. Scan Quality: What was the quality of the WB-MRI used for this read? *

	Good	Adequate	Poor
1. B 900	◯	◯	◯
2. ADC	◯	◯	◯
3. T1 sequences	◯	◯	◯

37. Please enter any specific comments you have on scan quality

38. Reader confirmation: My responses have been accurately reported on this CRF (enter 'yes' if in agreement with this statement) *

- ○ Yes
- ○ No

Supplementary S4 – MALIMAR Amendments

No. and Type of Amendment	Date approved	Brief Details of amendment
1. Non-substantial	25/06/2018	Protocol administrative updates
2. Non-substantial	15/01/2019	Communications to HVs
3. Non-substantial	19/03/2019	Update on scan numbers for protocol
4. Non-substantial	16/10/2019	Addition of ICHT site
5. Non-substantial	28/06/2019	Extension of project time-line and uplift in costs

SPIRIT Checklist for *Trials*

Complete this checklist by entering the page and line numbers where each of the items listed below can be found in your manuscript.

Your manuscript may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please state "n/a" and provide a short explanation. **Leaving an item blank or stating "n/a" without an explanation will lead to your manuscript being returned before review.**

Upload your completed checklist as an additional file when you submit to *Trials*. You must reference this additional file in the main text of your protocol submission. The completed SPIRIT figure must be included within the main body of the protocol text and can be downloaded here: http://www.spirit-statement.org/schedule-of-enrolment-interventions-and-assessments/

In your methods section, please state that you used the SPIRIT reporting guidelines, and cite them as:

Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin J, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. SPIRIT 2013 Explanation and Elaboration: Guidance for protocols of clinical trials. BMJ. 2013;346:e7586

	Reporting Item		Page and Line Number	Reason if not applicable
Administrative information				
Title	#1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	Page 1, line 1	
Trial registration	#2a	Trial identifier and registry name. If not yet registered, name of intended registry	Page 7, line 141	
Trial registration: data set	#2b	All items from the World Health Organization Trial Registration Data Set	n/a	Not a RCT
Protocol version	#3	Date and version identifier	Page 22, line 488	

Funding	#4	Sources and types of financial, material, and other support	Page 4, line 71	
Roles and responsibilities: contributorship	#5a	Names, affiliations, and roles of protocol contributors	Page 4, line 65	
Roles and responsibilities: sponsor contact information	#5b	Name and contact information for the trial sponsor	Page 4, line 86	
Roles and responsibilities: sponsor and funder	#5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	Page 4, line 86 Page 20, line 452	
Roles and responsibilities: committees	#5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	Page 20, line 440	
Introduction				
Background and rationale	#6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	Page 8, line 156	

Background and rationale: choice of comparators	#6b	Explanation for choice of comparators	Page 9, line 181 - 191	
Objectives	#7	Specific objectives or hypotheses	Page 9, line 191	
Trial design	#8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)	Page 9, line 199 Page 12 line 254	
Methods: Participants, interventions, and outcomes				
Study setting	#9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	Page 10, line 216	
Eligibility criteria	#10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	Page 10-11, line 221-228 and Table 1	
Interventions: description	#11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	Page 11, line 229	
Interventions: modifications	#11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving / worsening disease)	n/a	No modifications

Interventions: adherance	#11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return; laboratory tests)	n/a	No strategies or monitoring of adherance
Interventions: concomitant care	#11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	n/a	Use of scans only
Outcomes	#12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	Page 15, line 343	
Participant timeline	#13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Figure 1 and Table 3	Schedule and assessment of scans rather than participants
Sample size	#14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	Page 16, lines 366	
Recruitment	#15	Strategies for achieving adequate participant enrolment to reach target sample size	Page 12, line 271 Page 13, line 285	

Methods: Assignment of interventions (for controlled trials)				
Allocation: sequence generation	#16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	Page 10, lines 194-210	
Allocation concealment mechanism	#16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	Page 10, lines 200-210	
Allocation: implementation	#16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	Page 14, line 316 Page 15, line 327	
Blinding (masking)	#17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	Page 14, line 318 Page 21, line 470	
Blinding (masking): emergency unblinding	#17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	n/a	No blinding of intervention
Methods: Data collection, management, and analysis				

Data collection plan	#18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	Page 15, line 337	
Data collection plan: retention	#18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	n/a	No participant retention/follow-up
Data management	#19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	Page 15, line 340	
Statistics: outcomes	#20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	Page 17, line 391-428	
Statistics: additional analyses	#20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	n/a	Primary and secondary analysis included only

Statistics: analysis population and missing data	#20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	Page 19, line 421	
Methods: Monitoring				
Data monitoring: formal committee	#21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	Page 19 line 433 – line 450	
Data monitoring: interim analysis	#21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	Page18, line 396 Page 18, line 408	
Harms	#22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	Page 19, line 433	No adverse event reporting as not an interventional study on participants. Specified in manuscript.
Auditing	#23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	Page 19, line 438	
Ethics and dissemination				

Research ethics approval	#24	Plans for seeking research ethics committee / institutional review board (REC / IRB) approval	Page 20, line 452	
Protocol amendments	#25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC / IRBs, trial participants, trial registries, journals, regulators)	Page 20, line 456	
Consent or assent	#26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	Page 13, lines 272-279	
Consent or assent: ancillary studies	#26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	n/a	No ancillary studies
Confidentiality	#27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	Page 20, line 458-461 Page 22, line 484	
Declaration of interests	#28	Financial and other competing interests for principal investigators for the overall trial and each study site	Page 5, line 102	
Data access	#29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	Page 20, line 457-458	

Ancillary and post trial care	#30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	n/a	No participant level intervention
Dissemination policy: trial results	#31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	Page 20, lines 457-461	
Dissemination policy: authorship	#31b	Authorship eligibility guidelines and any intended use of professional writers	Page 20, lines 457-461	
Dissemination policy: reproducible research	#31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	Page 20, lines 457-461	
Appendices				
Informed consent materials	#32	Model consent form and other related documentation given to participants and authorised surrogates	Supplementary material S1a, S1b	
Biological specimens	#33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	n/a	No biological specimens collected

It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-