Appendix 1. METastasis Reporting and Data System for Prostate cancer baseline assessments

All suspicious lesions seen on whole-body magnetic resonance imaging (WB-MRI) are mapped on the baseline regional template form and baseline measurements should also be recorded on the measurement template form.

1.1. Baseline regional template form instructions

This allows for clear clinical documentation of the involvement over 14 anatomic regions:

Primary disease (1)	
Bone disease:	Skull (2)
	Cervical spine (3)
	Dorsal spine (4)
	Thorax (5)
	Lumbo-sacral spine (6)
	Pelvis (7)
	Extremities (8)
Visceral disease:	Pelvic lymph nodes (9)
	Retroperitoneal nodes (10)
	Lymph nodes, other (11)
	Lungs (12)
	Liver (13)
Other sites (14)	

After first deciding on whether there is evidence of involvement of each region, the reporting radiologist reports on the number of lesions for each bone and visceral region, on the baseline template reporting form.
These values provide ranges and descriptors for the potential values to aid in later comparisons. The number system is designed to reduce the time taken to quantify large numbers of lesions within regions.
Instructions

- For each region, involvement, lesion number, and morphologic characteristics should be recorded. Diffuse disease is recorded as a separate category.
- Note also the presence/absence of primary disease and assess local complications—rectal, bladder, and ureteric invasion.

1.1. METastasis Reporting and Data System for Prostate Cancer Baseline regional reporting template form				
Date	Region	Involved: Y or N or N/A	Lesion no.	Description
Local disease	Primary disease		–	Involvement of adjacent structures recorded here
Bone disease[c]	Skull			
	Cervical spine			
	Dorsal spine			
	Lumbo-sacral spine			
	Thorax[a]			
	Pelvis			
	Extremities[b]			
Visceral disease	Lymph nodes (pelvis)			
	Lymph nodes (retroperitoneum)			
	Lymph nodes (other)			
	Liver[e]			
	Lung[d]			
	Other sites			
Comments:				

N = no; N/A = not applicable; Y = yes.

[a] Thorax: ribs, sternum, scapula, clavicle.

[b] Extremities includes proximal humeri and femora.

[c] Bone lesions: 1; 2; 3–5; 6–10; ≥10; diffuse.

[d] Lungs: 1; 2; 3–5; 6–10; ≥10; diffuse.

[e] Liver: 1; 2; 3–5; 6–10; ≥10; diffuse.

1.2. Measurements template form instructions

- This template form is designed to collect objective dimensional measurements at baseline and on follow-up over time (mainly for clinical trials purposes).
- Lesion measurements should be undertaken on morphologic T1 weighted (W) and T2W images whenever possible. The large matrix size of diffuse weighted imaging (DWI) and apparent diffusion coefficient (ADC) maps results in magnification of dimensions of lesions. Partial volume effects on DWI can also result in uncertainty regarding lesion dimensions particularly in low-contrast situations.
- The data is collected separately for local, nodal, visceral, and bone disease.
- Measurements of bone lesions, lymph nodes, and soft tissue assessment should also be undertaken using the measurement template form.
- Linear dimensions cannot be obtained in the setting of diffuse disease.
- Only discrete lesions should be evaluated in their longest dimension for local soft tissue disease, visceral disease, and bones; short axis to be used for nodal disease (the plane of imaging will depend on the lesion location).
- Up to five bone, five nodes, and five soft tissue lesions should be measured (15 lesions maximum).
- A ≥1.5-cm threshold applies to all measured lesions (see preamble for exceptions).
- Only measurable disease as defined below can be included.
- Lesion locations including sequence and image numbers should be recorded to allow serial objective measurements on follow-up studies.

Local prostate disease

- Local disease is considered separately from nodal, bone, and visceral disease
- Local recurrence ≥1.5 cm is considered measurable.
- Local disease >1 cm and <1.5 cm is considered evaluable but not measurable unless high resolution; small field of view images are obtained when >1 cm can be considered measurable.

Nodes

- Record pelvic, retroperitoneal, and other nodes separately. Up to five nodes (all regions) are measured in total.
- Nodes ≥1.5 cm in short axis diameters are considered measureable. Nodes ≥ 1.0 cm and < 1.5 cm in short axis are considered pathologic but nontarget (nonmeasureable). Nodes <1.0 cm are considered nonpathologic.

Visceral lesions

- Visceral disease to be separately recorded and distinguished from nodal and bone disease.
- Up to five visceral lesions should be recorded in total, not more than two lesions per organ.
- Lesions ≥1.5 cm in longest dimensions are considered measureable and can be chosen as target and non-target lesions for response assessments.
- Lesions <1.5 cm are considered non measureable.

Bones

- Up to five bone lesions should be recorded in total with at least one lesion (≥1.5 cm) in appendicular skeleton (when present).
- Where possible, not more than two lesions per bone (each hemi-pelvis counted separately).
- Lesions should be measured on T1W images in the longest dimension.
- Lesions ≥1.5 cm is considered to be measurable (for exception see preamble).
- On follow-up studies, when focal lesions become diffuse and involve the entire bone marrow without extra-osseous soft-tissue, then measures of bone edge-to-edge should be undertaken.
- Signal intensity extent is a subjective assessment referring to extent (not intensity) of abnormalities consistent with cancer compared with immediate prior study, taking into account the caveats already described (not required at baseline).
- *Diffuse bone disease*
    - Diffuse bone disease cannot be recorded for size measurements. Indicate *diffuse* in the size column.
    - Diffuse bone disease should have ADC measurements and their location noted (to include the entire involved bone, avoiding the outer cortical margin, neural foramina, and intraosseous vessels).
    - Two bones representative of diffuse bone involvement should be chosen. Suggested locations include lower lumbar spinal vertebrae and posterior iliac bones (if free of artefacts).
    - The anatomical sites for the measurements should be recoded.
- *ADC measurements in bone disease*
    - ADC measurements should only be obtained from bones when there is sufficient signal intensity detected on b-value images (including b0); otherwise the ADC values will be erroneous, reflecting only the noise in the images.
    - The absence of tissue signal on high b800–1000 b-value images does not exclude the tissue from ADC measurements. Low and intermediate b-value images should be chosen instead region of interest measurements. See Appendix Figure 3 on the effectiveness of this approach to obtain more representative post therapy ADC values.

1.2. METastasis Reporting and Data System for Prostate Cancer measurements template form (baseline and follow-up studies)

Bone metastases measurements	Lesion (location)	Date
Soft tissue Y/N	***Lesion 1*** *Sequence/image*	
Size (mm/cm) for diffuse infiltration indicate		
ADC mean (in $10^{-6}mm^2/s$) diffuse disease indicate location		
Signal intensity extent (increased, decreased, stable) (follow-up studies)		
Soft tissue Y/N	***Lesion 2*** *Sequence/image*	
Size (mm/cm) for diffuse infiltration indicate		
ADC mean (in $10^{-6}mm^2/s$) diffuse disease indicate location		
Signal intensity extent (increased, decreased, stable) (follow-up studies)		
Soft tissue Y/N	***Lesion 3*** *Sequence/image*	
Size (mm/cm) for diffuse infiltration indicate		
ADC mean (in 10^{-6} mm^2/s) diffuse disease indicate location		
Signal intensity extent (increased, decreased, stable) (follow-up studies)		
Soft tissue Y/N	***Lesion 4*** *Sequence/image*	
Size (mm/cm) for diffuse infiltration indicate		
ADC mean (in $10^{-6}mm^2/s$) diffuse disease indicate location		
Signal intensity extent (increased, decreased, stable) (follow-up studies)		
Soft tissue Y/N	***Lesion 5*** *Sequence/image*	
Size (mm/cm) for diffuse infiltration indicate		
ADC mean (in $10^{-6}mm^2/s$) diffuse disease indicate location		
Signal intensity extent (increased, decreased, stable) (follow-up studies)		

Soft tissue metastases measurements	Image reference	*Baseline/visit x (date)*
Target[a]		
1 *Description including location*	Sequence/image	Size (mm/cm)
2 *Description including location*	Sequence/image	Size (mm/cm)
3 *Description including location*	Sequence/image	Size (mm/cm)
4	Sequence/image	Size (mm/cm)
5	Sequence/image	Size (mm/cm)
6		
7		
8		
9		
10		
Nontarget lesions[b] (list) *Description including location*	Follow-up visits Presence/absence Increase/stable/decrease/resolved	
Nonmeasureable lesions[c] (list) *Description including location*	Follow-up visits Presence/absence Increase/stable/decrease/resolved	
New lesions (list) *Description including location* *Nonbaseline studies only*	Sequence/image	
Local disease *Description*	Sequence/image	Size (mm/cm)

ADC = apparent diffusion coefficient; N = no; Y = yes.

[a] Include up to five lymph nodes and five visceral lesions. Not more two lesions per organ. No bone measurements (use prior form).
[b] Nontarget: measurable disease that are not specifically measured.
[c] Nonmeasureable lesions: lesions/disease not qualified for measurements.

Appendix 2. METastasis Reporting and Data System for Prostate Cancer follow-up assessments

Follow-up assessments methods instructions

Follow-up assessments use the measurements template form (Appendix 1.2) and the follow-up regional assessment template forms 1 and 2 (Appendix 2.1 and 2.2).

Measurement template form (Appendix 1.2)

Detailed instructions are given in Appendix 1.

Regional response assessment template forms 2.1 and 2.2

- In order to complete template forms 2.1 and 2.2 the changes seen within each region are categorised according to the response assessment categories (RACs) shown in Appendix Table 2.
- The RACs have been designed to enable regional assessments of treatment response to take place in a reproducible and semiquantifiable way.
- RAC for regional assessments use a sliding scale. These RACs are to be applied to both soft tissues and to bone regions using the criteria given in Table 3 of the main paper.
- There are five potential RACs on a sliding scale (highly likely to be responding, likely to be responding, stable, likely to be progressing, and highly likely to be progressing).
- The Prostate Cancer Clinical Trials Working Group modifications of Response Criteria in Solid Tumours (RECIST) v1.1 are built into the RACs for soft tissue assessments (Table 3) [1], as part of the definitions for RAC 1 and RAC 5. The definition of RAC 1 and RAC 5 for bone regions is based on the work of Lecouvet et al [2].

Instructions

Template forms 2.1 and 2.2 pictorially represent the same 14 regions assessed at baseline. The methodology for completing these follow-up regional template forms is:

- Each of the 14 regions is separately assessed for evidence of disease involvement. If disease is present within a region then further assessment is made of the treatment response (within this region only) by comparison with the baseline scan, as described below.
- A primary RAC value (1–5) is assigned to the region based on dominant pattern of response within the region. This is defined as the response shown by more than half of the lesions within the region. RAC criteria are defined in Table 3 in the main paper.
- A secondary RAC value (1–5) is assigned to the region to illustrate the second most frequent pattern of response seen within the region in question.
- A tertiary RAC value (4–5) is assigned to the region to illustrate that there is evidence of progressing disease (ie, RAC 4–5) but that this is neither the

dominant nor secondary pattern of response within the region (ie, not captured by the primary or secondary RAC values).

- For a single lesion per region only the primary number category is assessed (see Appendix Fig. 4 for illustration of usage).
- Regions with multiple lesions all with the same pattern of response will have the same RAC value assigned as both the primary and secondary RACs.
- When >1 lesion is present and equal numbers of lesions are category RAC 4/5 as RAC 1/2/3, then the primary pattern allocation is reserved for RAC 4/5.
- Similarly, when >1 lesion is present and equal numbers of lesions are category RAC 2 as RAC 3, then the primary pattern allocation is reserved for RAC 3 (the higher category).

The above methodology allows for the capture of discordant treatment responses within anatomical regions. For example, a region given a score of 1 + 3 shows a dominant pattern that is highly suggestive of response, with a secondary pattern showing no change. A worked-up example of documentation of regional responses is given in Appendix Figure 4.

2.1. METastasis Reporting and Data System for Prostate Cancer regional response assessment template form 1: Skeletal

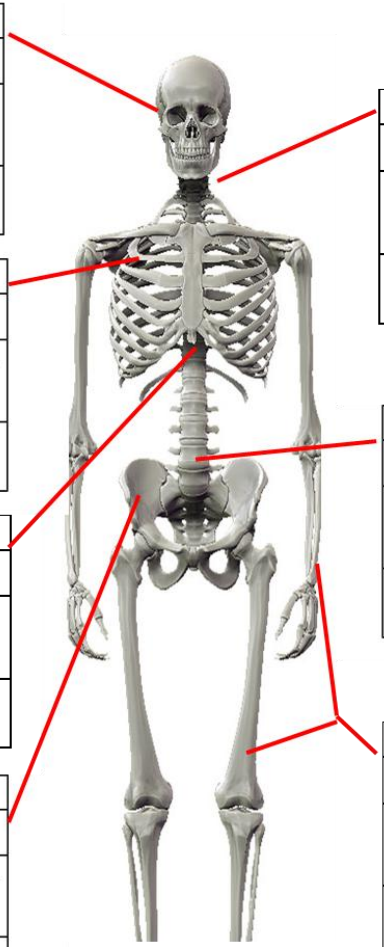
Pt identifier: ……. Scan date: …./…./…. Scan number: ……. Reporter: ……..

SKULL			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

RIBS AND SHOULDER GIRDLE			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

THORACIC SPINE			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

PELVIS			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			



CERVICAL SPINE			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

LUMBO-SACRAL SPINE			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

EXTREMITIES (Upper and Lower Limbs)			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

2.2. METastasis Reporting and Data System for Prostate Cancer regional response assessment template form 2: Soft tissue

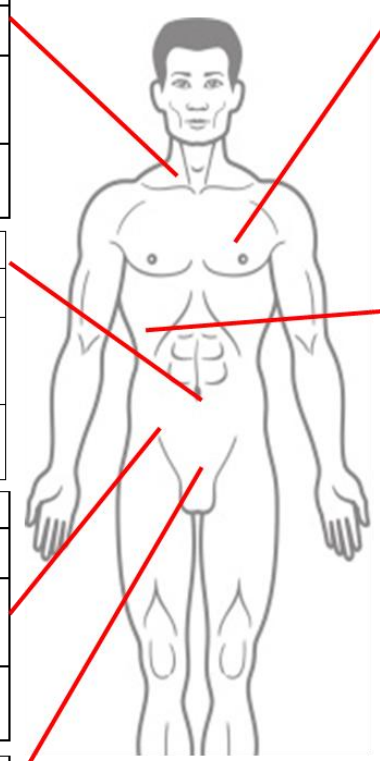
Pt identifier: ..... Scan date: ..../..../.... Scan number: ..... Reporter: .....

OTHER METASTATIC LYMPH NODES			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

PARA-AORTIC LYMPH NODES			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

PELVIC LYMPH NODES			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

PRIMARY DISEASE			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			



LUNGS			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

LIVER			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

OTHER SITES			
Involved	Yes / No / Not assessable		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			

Appendix 3. METastasis Reporting and Data System for Prostate Cancer overall assessments

Instructions

The status of the primary disease, nodes, viscera, and bone disease should be recorded separately using the overall response assessment template form.

Unlike regional response assessments, overall response for the primary tumour, nodal, and visceral disease should be categorical, thus following established guidelines (Prostate Cancer Clinical Trials Working Group modifications of RECIST v1.1) [1]. The following categories should be assigned: complete response, partial response, stable disease, progressive disease, and discordant.

Note that progression assignments for soft tissues are based on measurements and should be from baseline or treatment induced nadir whichever is lower. Other progression assignments are as per RECIST v1.1 (eg, new disease; Table 3) [3].

In contradistinction, the overall response of bone disease should be categorised on a scale of 1 to 5 indicating the likely overall response category: (1) highly likely to be responding, (2) likely to be responding, (3) stable, (4) likely to be progressing, and (5) highly likely to be progressing. Bone disease uses the criteria given Table 3.

The type of progression (new disease versus growth of existing lesions) should be separately recorded; the location of progression should be obtainable from the regional response assessment template forms 2.1 and 2.2

Discordant, progressive disease should also be separately reported for primary, nodal, viscera, and bone; evaluation of regional discordant responses on forms 2.1 and 2.2 will enable the specific identification of the anatomic sites of mixed responses.

Discordance indicates the presence of progressing bone/soft tissue disease, not meeting definite progression criteria in the primary category, that is, when the majority of disease is stable or responding. In each discordant case, indicate whether discordance is a secondary (ie, major discordance) or tertiary (ie, minor discordance) assessment (Appendix Table 2).

<table>
<tr><td colspan="5">3.0. METastasis Reporting and Data System for Prostate Cancer overall response assessment categories</td><td colspan="2">Date</td><td colspan="2">Patient label</td></tr>
<tr><td colspan="9">Nodal disease</td></tr>
<tr><td rowspan="2">No disease</td><td rowspan="2">Complete response</td><td rowspan="2">Response</td><td rowspan="2">Stable</td><td colspan="2">Progression</td><td colspan="4" rowspan="2"></td></tr>
<tr><td>New</td><td>Growth</td></tr>
<tr><td>Discordant [yes/no]</td><td>–</td><td>Responding-discordant Major/minor</td><td>Stable-discordant Major/minor</td><td colspan="2">–</td><td colspan="4"></td></tr>
<tr><td colspan="9">Visceral disease</td></tr>
<tr><td rowspan="2">No disease</td><td rowspan="2">Complete response</td><td rowspan="2">Response</td><td rowspan="2">Stable</td><td colspan="2">Progression</td><td colspan="4" rowspan="2"></td></tr>
<tr><td>New</td><td>Growth</td></tr>
<tr><td>Discordant [yes/no]</td><td>–</td><td>Responding-discordant Major/minor</td><td>Stable-discordant Major/minor</td><td colspan="2">–</td><td colspan="4"></td></tr>
<tr><td colspan="9">Local disease</td></tr>
<tr><td rowspan="2">No disease</td><td rowspan="2">Complete response</td><td rowspan="2">Response</td><td rowspan="2">Stable</td><td colspan="2">Progression</td><td colspan="4" rowspan="2"></td></tr>
<tr><td>New</td><td>Growth</td></tr>
<tr><td>Discordant [yes/no]</td><td>–</td><td>Responding-discordant Major/minor</td><td>Stable-discordant Major/minor</td><td colspan="2">–</td><td colspan="4"></td></tr>
<tr><td colspan="9">Bone disease</td></tr>
<tr><td rowspan="2">No disease</td><td>Response</td><td>Likely Responding</td><td rowspan="2">Stable</td><td colspan="2" rowspan="2">Likely progressing</td><td colspan="2">Definite progression</td><td colspan="2" rowspan="2"></td></tr>
<tr><td colspan="2">New</td><td>Growth</td></tr>
<tr><td>Discordant [yes/no]</td><td>Response-discordant Major/minor</td><td>Responding-discordant Major/minor</td><td>Stable-discordant Major/minor</td><td colspan="2">–</td><td colspan="4"></td></tr>
</table>

Appendix 4. Machine set-up, quality assurance, and quality control

The radiologic literature and machine manufacturer recommendations should be used by radiologists/technologists/physicists for machine set-up and to adapt imaging protocols for the level of machine software [4–10].

4.1. Machine set-up for clinical care and research

WB-MRI performed at the field strength of 1.5T has become the established platform due to its robustness [4] and widespread availability; however, excellent results can also be obtained with increased signal-to-noise ratio (SNR) at 3T, on many modern systems, including wide bore systems [11].

In the context of evaluating tumour response to treatments, whenever possible repeat examinations should be performed on the same machine (type and software version) during follow-up studies. We do not recommend same patient measurements at different field strengths, even from a single vendor with identical sequence software versions. See **Appendix Figure 1** for an example of marked image variations due to changes in magnetic field strength.

There is unanimous consensus regarding the positive benefits of using surface coils to maximise the SNR of images; signal reception with machine integrated body coil is thus not recommended, although the latter may be used according to the clinical situation, local expertise, and the equipment available or for patient comfort. In these cases, the SNR loss resulting from a lack of surface coils will need to be countered by sequence adjustments such as increasing acquisition pixel size, increasing the number of averages, and reduced pixel bandwidth, which will lead to penalties, including longer scan times, and reduced spatial resolutions.

MR clinical scientists working with experienced radiographers/technologists should undertake vendor specific WB-MRI set-ups. Sequences adjustments and optimisation of measurement protocols should be undertaken using both volunteers and test objects [12,13].

Following agreements with responsible radiologists, the measurement protocol parameters including length of torso coverage, in-plane field of view (FOV), image matrix, corresponding voxel sizes, slice thickness, fat-suppression methods for diffusion sequences should be determined and fixed to enable intra- and inter-patient comparisons within study cohorts. Care should be taken when employing vendor provided workflow optimisation software, which can unexpectedly change sequence parameters and image matrices without user notifications.

Since effective fat suppression is a prerequisite for water ADC assessments in bone marrow both olefinic and aliphatic compounds require suppression, this is ideally realized using inversion recovery methods (short tau inversion recovery [STIR]). The effectiveness of fat suppression should be evaluated in normal volunteers with a range of body mass index and suitable test individuals as detailed below under the quality assurance section.

For research studies, it is additionally necessary to measure and document repeatability and reproducibility of diffusivity measures (including ADC) during trial set-up, using volunteers to evaluate various body tissues at each bed position in the midline and off-centre.

4.2. Machine quality assurance

The quality assurance (QA) program should as far as possible be under the supervision of clinical scientists/medical physicists for research studies but maybe undertaken by trained experienced technologists/radiographers for routine clinical studies.

Many clinical departments do not perform routine QA measurements, partly as there are no established tolerances to indicate at which point remedial action is to be taken. However, a general QA program should be established according to manufacturer recommendations, and extended to include DWI when quantitative ADC assessments measurements are being used for lesion detection, characterization, and for response assessments [14].

Initial set-up of quantitative WB-DWI measurements should include assessments of uniformity of ADC over large FOV in all three axes, evaluations of B0 distortions and measures of ghosting resulting from, echo-planar imaging readout, parallel imaging, and eddy currents. Test objects should be employed to optimise fat suppression over large FOVs should be used [13]. This applies equally to clinical and research uses.

Detailed DWI QA measurements should be performed and documented at least monthly for research studies and more than 6 monthly for clinical applications.

Consideration should be given to repeating routine and DWI QA measurements after repairs and routine maintenance, after cryogen fills, and whenever there are adjustments to the scanner default parameters. Extra QA checks are needed after software or hardware upgrades (Appendix Fig. 1).

Deviations in established measurement performance should be reported to a clinical scientist/medical physicist for evaluation and remedial action.

To enable comparisons of diffusion weighted images between machines, QA measurements of b-value signal intensities, SNR and ADC measurements should be undertaken using biologic tissues with low variance in diffusion properties such as the brain, or test objects made of bioequivalent materials.

Suitable test materials for ADC assessment include polymerpolyvinylpyrrolidone solutions or sucrose solutions [15] and iced-water phantoms, although designed for smaller FOV imaging it may be adapted for these purposes [13,16]. Test objects using corn oil [16] for fat fraction assessment and fat suppression may be also used.

Detailed recommendations on QA procedures for quantitative WB-DWI can be found on the UK Quantitative WB-DWI Technical Workgroup 2016 website (https://sites.google.com/site/wbadcconsensus/home).

4.3. Image quality control

It is important to emphasise the need for consistent patient preparation, patient positioning, and scanning procedures in order to obtain consistent reproducible imaging to allow comparisons between studies.

Patients should be asked to lie on the MR table in a head first, supine position, with their head resting in the appropriate head coil and their arms by their sides (where possible). Generally, if only the axial skeleton is to be scanned a leg-rest/knee support should be used as it provides patient comfort and improves patient compliance.

In standard WB-MRI, surface receiver coils should be placed on top of the patient ensuring the length of head-feet coverage. There should be no gaps in the coil coverage between these positions. The coils used should be site specific and should be the same for all patients.

Attention to patient comfort and adequate pain control are essential, especially as comprehensive WB-MRI scans can take between 45 min and 60 min to complete (depending on machine performance and anatomic imaging coverage required).

Taking these factors into consideration, the supervising radiologist/technologist/clinical scientist should optimise imaging protocols in order to obtain the best and most consistent image quality possible on the MRI scanner(s) used at their institution/centre. Imaging sequences for WB-MRI are detailed in Appendix 5 and in Table 2 of the main paper.

Technologists performing the examination and/or supervising radiologists should review image quality at the time of acquisition. They should be trained to detect the most common image artefacts and the approaches to correct these problems. Consequently, if image quality of a measurement is compromised due to patient motion or for another reason, remedial measures may be undertaken.

Additional regional imaging should be undertaken to evaluate disease sites/anatomic regions as needed (for local tumour recurrence, dedicated small FOV images to evaluate spinal cord compression or neural foramina compromise, contrast enhancement for brain evaluations, etc.).

Appendix 5. Sequence specifications

WB-MRI studies should always include combinations of imaging sequences (T1W, T2W, STIR) and DWI. Bone marrow fat and water imaging using Dixon techniques, dedicated anatomic images of body parts and contrast medium enhancement maybe used as indicated in Table 2 of the main paper**.**

T1W images are useful for the detection of visceral, nodal, and skeletal metastases. A combination of two-dimensional spin-echo and gradient-recalled echo (GRE) techniques can be utilised as detailed in Table 2 of the main paper. In order to speed up spin-echo sequences, a fast or turbo-spin echo sequence with an echo-train length of 3–5 may be employed. When GRE sequences are used, we recommend the Dixon technique [17,18] because it enables the calculation of bone marrow fat fraction [19,20]. As an alternative to the two-dimensional techniques above, a three-dimensional turbo-spin echo T1W sequence can be acquired instead if machine performance allows [10].

T2W images are used to depict soft tissue anatomy, to assess visceral organs including the lungs, liver, kidneys, rectum, and bladder for local and distant tumour spread and to assess complications induced by the primary tumour. Fat suppression of T2W images is not routinely recommended; low b-value DWI (b50–100) can be used as an anatomic, fat-suppressed T2W imaging substitute if required (to detect bone or soft tissue oedema, for the detection of free fluid, or for coregistration of images).

Both T2W and STIR sequences are helpful for assessing the spinal canal, for disease encroachment, and spinal cord compression detection. Traditionally, whole body STIR imaging was undertaken as part of WB-MRI evaluations, but these are no longer recommended. Rarely, when the quality of the spinal STIR images are suboptimal, a T2W fat suppressed sequence can be used as an alternative, although uneven fat suppression may occur. Alternate methods of fat suppression of the spine T2W images include Dixon sequences.

DWI of the whole body are considered essential for WB-MRI evaluations. Spin-echo echo-planar imaging sequences at multiple stations are required, to cover the torso from the skull base to the mid-thighs in the axial plane. When needed, imaging can be extended cranially to include the entire skull. Typical parameters for the major manufacturers are given in

Appendix Table 1.

There are a number of factors that need to be considered when choosing the b-values of DWIs. There is a need to retain an outline of the body to enable registration of images to anatomy sequences and for subsequent comparisons. Image artefacts (eg, distortion, susceptibility, and poor fat suppression) need to be minimised whilst maximising tumour/bone marrow contrast to background ratio. There is a need to minimise perfusion and maximise sensitivity to bone marrow cellular content. Since high b-value images are also used to evaluate normal and pathologic soft tissues (liver, nodes, etc.), intermediate b-values may also be needed. Finally, there is an absolute need to obtain reliable estimates of ADC values.

Taking these considerations into account, two b-values are recommended as a minimum (b50–100 s/mm^2 and b800–1000 s/mm^2), with an intervening b500–600-s/mm^2 image set obtained optionally, to stabilise ADC calculations and to provide improved SNR for the assessments of bone, liver, and nodal disease.

Acquisition of ultra-high b-value images (≥1000 s/mm^2) is discouraged, due to greater image distortions and the need for longer echo-times resulting in poor SNR. However, ultra-high b-value images may be extrapolated from the *source* DWI images without scanner time penalty, to facilitate image segmentation of bone metastases [21].

ADC map calculations should as far as possible and not include b0 images (even if b0 images can be obtainable without incurring time penalties). ADC maps should be performed by mono-exponential fitting of the signal intensities of the acquired b-value images. Judicious use of a background image filter maybe helpful; too strong a noise filter applied before estimating the ADC map may result in a loss of anatomical detail; the outline of the body torso needs to be retained on ADC images.

Bone marrow fat imaging

Dixon imaging is a chemical-shift technique that relies on phase differences between the resonance frequency of fat and water, to separate out tissue fat and water fractions. This is done by acquiring images at several echo times, each of which has a different resulting signal, attributable to fat and water fractions and the phase differences between these

fractions. Thus, four sets of images are produced (in-phase, opposed phase, fat only, and nonfat [water] only images), which can be used to calculate fat-fraction images (F%).

We recommend using a 2-point (echo) Dixon technique recognising that this method does not correct for T2* effects, which can be done by utilising a multi-echo technique [22,23]. Although multipoint, proton density weighted Dixon sequences yield the most accurate F% images [19], the source images have poor image contrast and serve little diagnostic purpose, and so are not recommended. Instead, we suggest the use of a 2-point, T1W, Dixon sequence for the whole body, in the knowledge that there is a likely overestimate of measured bone marrow F%.

Dixon T1-weighted images enable the detection of bone marrow and visceral metastases because marrow fat and liver are of higher signal intensity than metastases. Opposed phase images can be useful for bone lesion characterisation. Since the bone marrow metastasis detection performance of T1W GRE sequences is slightly worse that (turbo) spin echo sequences [17], we suggest that relative F% images are always reconstructed and evaluated with DWI, for the presence of disease.

Contrast medium enhancement is not essential for WB-MRI when the aim is to obtain information on bone and nodal disease. Its value lies in evaluating the prostate gland for local tumour recurrence, and liver and brain for metastatic involvement.

Appendix 6. Suggested indications for WB-MRI in prostate cancer

- High-risk prostate cancer (Gleason ≥ 8 and/or prostate-specific antigen [PSA] ≥ 20 ng/ml) at presentation [24]. The core WB-MRI protocol may be performed to substitute for the combined use of bone scan and a body computed tomography (CT) scan [8]. A chest CT scan may still be required if there are anaplastic features histologically.
- Equivocal bone scans in newly diagnosed patients to clarify the nature or extent of abnormalities detected. Here a dedicated MRI scan targeted to an abnormality detected by another modality could also suffice.
- Biochemical recurrence in patients at high risk of developing metastases with or without a prior negative CT and bone scan (various thresholds for higher sensitivity imaging including positron emission tomography/CT are suggested in literature including PSA >2 ng/ml (postprostatectomy) or >5 ng/ml (postradiotherapy); PSA doubling time <6 mo; and after PSA doubling following prior negative imaging [25,26]. Here the aim is to detect oligo-metastatic disease (see below); the core WB-MRI protocol should be sufficient.
- Nonmetastatic castration-resistant prostate cancer with or without a prior negative CT and bone scan (various thresholds for higher sensitivity imaging are suggested in the literature including PSA > 2 ng/ml or PSA doubling following prior negative imaging [25,26]). Here the aim is to detect oligo-metastatic disease (see below); the core WB-MRI protocol should be sufficient.
- Detecting oligo-metastatic disease when radical therapy to metastatic disease may be an option, with the aim of postponing the onset of systemic treatments [27,28]; the core WB-MRI protocol should be sufficient.
- Rule out occult distant metastases in locally recurrent disease when salvage local therapy is being considered (confirm nonmetatastic disease); the core WB-MRI protocol should be sufficient.
- Known extensive spinal bone metastases on bone or CT scan to detect clinically occult spinal cord compression (spinal examination only will suffice). A

comprehensive WB-MRI protocol should be used as the designated follow-up method.

- To detect presence and volume of metastatic disease in castrate naïve prostate cancer being considered for upfront combination hormonal therapy and docetaxel. Here a comprehensive WB-MRI protocol is recommended as WB-MRI is the preferred follow-up method.
- To monitor response in metastatic patients with bone disease (particularly PSA non- or oligo-secretory disease); baseline and follow-ups scans needed for each therapy instituted. Here a comprehensive WB-MRI protocol is recommended as WB-MRI is the preferred follow-up method.
- Patients deemed to have *anaplastic or neuro-endocrine features* on clinical/biochemical/imaging grounds (including exclusive visceral or predominantly lytic bone metastases, bulky tumour masses, those with low PSA levels relative to tumour burden, and those who have short responses to primary androgen deprivation therapy) [29], are to be treated with chemotherapy. Here a comprehensive WB-MRI protocol is recommended as WB-MRI is the preferred follow-up method.
- To confirm bone predominant disease in symptomatic patients when considering radium-223 treatment. Here a comprehensive WB-MRI protocol is recommended because visceral or bulky nodal disease are therapy contraindications [30].

References

[1] Scher HI, Morris MJ, Stadler WM, et al. Trial design and objectives for castration-resistant prostate cancer: Updated recommendations from the Prostate Cancer Clinical Trials Working Group 3. J Clin Oncol 2016:1–38. http://dx.doi.org/10.1200/JCO.2015.64.2702

[2] Lecouvet FE, Larbi A, Pasoglou V, , et al. MRI for response assessment in metastatic bone disease. Eur Radiol 2013;23:1986–97. http://dx.doi.org/10.1007/s00330-013-2792-3

[3] Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228–47. http://dx.doi.org/10.1016/j.ejca.2008.10.026

[4] Koh D-M, Blackledge M, Padhani AR, et al. Whole-body diffusion-weighted MRI: tips, tricks, and pitfalls. AJR Am J Roentgenol 2012;199:252–62. http://dx.doi.org/10.2214/AJR.11.7866

[5] Padhani AR, Koh D-M, Collins DJ. Whole-body diffusion-weighted MR imaging in cancer: current status and research directions. Radiology 2011;261:700–18. http://dx.doi.org/10.1148/radiol.11110474

[6] Kwee TC, van Ufford HM, Beek FJ, et al. Whole-body MRI, including diffusion-weighted imaging, for the initial staging of malignant lymphoma: comparison to computed tomography. Invest Radiol 2009;44:683–90. http://dx.doi.org/10.1097/RLI.0b013e3181afbb36

[7] Michielsen K, Vergote I, Op de Beeck K, et al. Whole-body MRI with diffusion-weighted sequence for staging of patients with suspected ovarian cancer: a clinical feasibility study in comparison to CT and FDG-PET/CT. Eur Radiol 2014;24:889–901. http://dx.doi.org/10.1007/s00330-013-3083-8

[8] Lecouvet FE, El Mouedden J, Collette L, et al. Can whole-body magnetic resonance imaging with diffusion-weighted imaging replace Tc 99m bone scanning and computed tomography for single-step detection of metastases in patients with high-risk prostate cancer? Eur Urol 2012;62:68–75. http://dx.doi.org/10.1016/j.eururo.2012.02.020

[9] Lecouvet FE. Whole-body MR imaging: Musculoskeletal applications. Radiology 2016;279:345–65. http://dx.doi.org/10.1148/radiol.2016142084

[10] Pasoglou V, Michoux N, Peeters F, et al. Whole-body 3D T1-weighted MR imaging in patients with prostate cancer: feasibility and evaluation in screening for metastatic disease. Radiology 2015;275:155–66. http://dx.doi.org/10.1148/radiol.14141242

[11] Koh D-M, Lee J-M, Bittencourt LK, Blackledge M, Collins DJ. Body diffusion-weighted MR imaging in oncology: Imaging at 3 T. Magn Reson Imaging Clin N Am 2016;24:31–44. http://dx.doi.org/10.1016/j.mric.2015.08.007

[12] Winfield JM, Collins DJ, Priest AN, et al. A framework for optimization of diffusion-weighted MRI protocols for large field-of-view abdominal-pelvic imaging in multicenter studies. Med Phys 2016;43:95–110. http://dx.doi.org/10.1118/1.4937789

[13] Winfield JM, Douglas NHM, Desouza NM, Collins DJ. Phantom for assessment of fat suppression in large field-of-view diffusion-weighted magnetic resonance imaging. Phys Med Biol 2014;59:2235–48. http://dx.doi.org/10.1088/0031-9155/59/9/2235

[14] Delakis I, Moore EM, Leach MO, De Wilde JP. Developing a quality control protocol for diffusion imaging on a clinical MRI system. Phys Med Biol 2004;49:1409–22

[15] Chenevert TL, Galbán CJ, Ivancevic MK, et al. Diffusion coefficient measurement using a temperature-controlled fluid for quality control in multicenter studies. J Magn Reson Imaging 2011;34:983–7. http://dx.doi.org/10.1002/jmri.22363

[16] Malkyarenko DI, Chenevert TL. Practical estimate of gradient nonlinearity for implementation of apparent diffusion coefficient bias correction. J Magn Reson Imaging 2014;40:1487–95. http://dx.doi.org/10.1016/j.drudis.2011.09.009

[17] Costelloe CM, Madewell JE, Kundra V, Harrell RK, Bassett RL, Ma J. Conspicuity of bone metastases on fast Dixon-based multisequence whole-body MRI: Clinical utility per sequence. Magn Reson Imaging 2013;31:669–75. http://dx.doi.org/10.1016/j.mri.2012.10.017

[18] Costelloe CM, Kundra V, Ma J, et al. Fast Dixon whole-body MRI for detecting distant cancer metastasis: a preliminary clinical study. J Magn Reson Imaging 2012;35:399–408. http://dx.doi.org/10.1002/jmri.22815

[19] Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration. J Magn Reson Imaging 2012;36:1011–4. http://dx.doi.org/10.1002/jmri.23741

[20] Ballon D, Watts R, Dyke JP, Lis E, Morris MJ, Scher HI, et al. Imaging therapeutic response in human bone marrow using rapid whole-body MRI. Magn Reson Med 2004;52:1234–8. http://dx.doi.org/10.1002/mrm.20291

[21] Blackledge MD, Collins DJ, Tunariu N, et al. Assessment of treatment response by total tumor volume and global apparent diffusion coefficient using diffusion-weighted MRI in patients with metastatic bone disease: a feasibility study. PLoS One 2014;9:e91779. http://dx.doi.org/10.1371/journal.pone.0091779

[22] Ballon D, Jakubowski AA, Tulipano PK, et al. Quantitative assessment of bone marrow hematopoiesis using parametric magnetic resonance imaging. Magn Reson Med 1998;39:789–800.

[23] Kühn J-P, Evert M, Friedrich N, et al. Noninvasive quantification of hepatic fat content using three-echo dixon magnetic resonance imaging with correction for T2* relaxation effects. Invest Radiol 2011;46:783–9. http://dx.doi.org/10.1097/RLI.0b013e31822b124c

[24] Bjurlin MA, Rosenkrantz AB, Beltran LS, Raad RA, Taneja SS. Imaging and evaluation of patients with high-risk prostate cancer. Nat Rev Urol 2015:1–12. http://dx.doi.org/10.1038/nrurol.2015.242

[25] Crawford ED, Stone NN, Yu EY, et al. Challenges and recommendations for early identification of metastatic disease in prostate cancer. Urology 2014;83:664–9. http://dx.doi.org/10.1016/j.urology.2013.10.026

[26] Rozet F, Roumeguère T, Spahn M, Beyersdorff D, Hammerer P. Non-metastatic castrate-resistant prostate cancer: a call for improved guidance on clinical management. World J Urol 2016. http://dx.doi.org/10.1007/s00345-016-1803-9

[27] Azzam G, Lanciano R, Arrigo S, et al. SBRT: An opportunity to improve quality of life for oligometastatic prostate cancer. Front Oncol 2015;5:101. http://dx.doi.org/10.3389/fonc.2015.00101

[28] Ost P, Jereczek-Fossa BA, AS NV, et al. Progression-free survival following stereotactic body radiotherapy for oligometastatic prostate cancer treatment-naive recurrence: A Multi-institutional analysis. Eur Urol 2016;69:9–12. http://dx.doi.org/10.1016/j.eururo.2015.07.004

[29] Aparicio AM, Harzstark AL, Corn PG, et al. Platinum-based chemotherapy for variant castrate-resistant prostate cancer. Clin Cancer Res 2013;19:3621–30. http://dx.doi.org/10.1158/1078-0432.CCR-12-3791

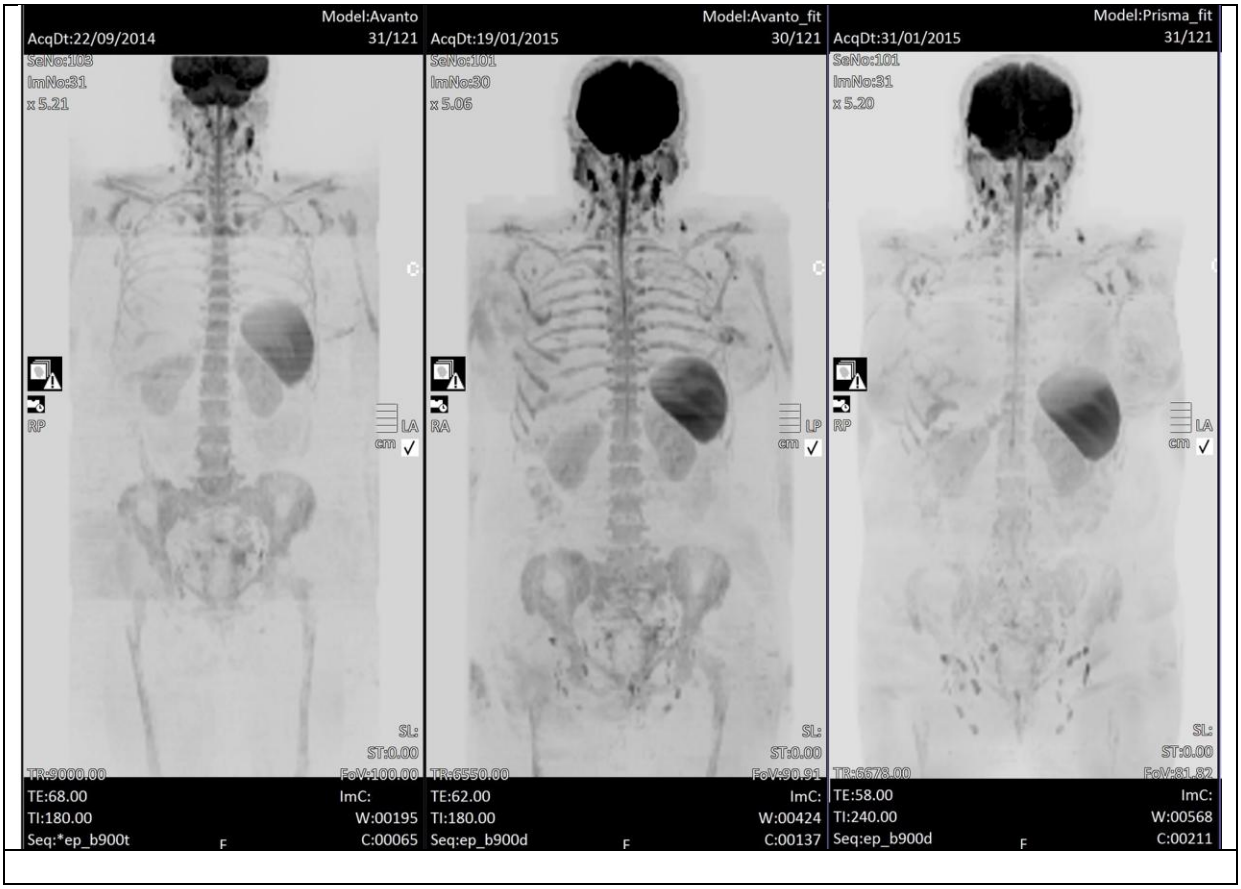
[30] Parker C, Nilsson S, Heinrich D, et al. Alpha emitter radium-223 and survival in metastatic prostate cancer. N Engl J Med 2013;369:213–23. http://dx.doi.org/10.1056/NEJMoa1213755

Appendix Table 1 – Whole body diffusion weighted magnetic resonance imaging sequence parameters			
	Siemens' 1.5T Avanto system	Philips 1.5T Intera system	GE 1.5T Signa system
Imaging plane	Axial	Axial	Axial
Field of view (cm)	400	400	400
Matrix size	128–150	128	128
Repetition time (ms)	6000–8000	8322	6625
Echo time (ms)	Min	Min	Min
Parallel imaging factor	2	2	2
No. of signals averaged for high b-value images	4–5 (reduce for low b-value if available)	4 (b = 50) 12 (b = 800–1000)	4 (reduce for low b-value if available)
Section thickness (mm)	5–7 contiguous	5–7 contiguous	5–7 contiguous
Direction of motion probing gradients	3 scan trace	Trace	All
Receiver bandwidth	1800 Hz/pixel	7.757 (water-fat shift/pixil)	976 Hz/pixel
Fat suppression	STIR (TI = 180 ms)	STIR (TI = 180 ms)	STIR (TI = 170 ms)
b-values (s/mm^2)	Typically 50–100 and 800–1000	Typically 50–100 and 800–1000	Typically 50–100 and 800–1000
Acquisition time per station	4 min 30 s	4 min 2 s	4 min

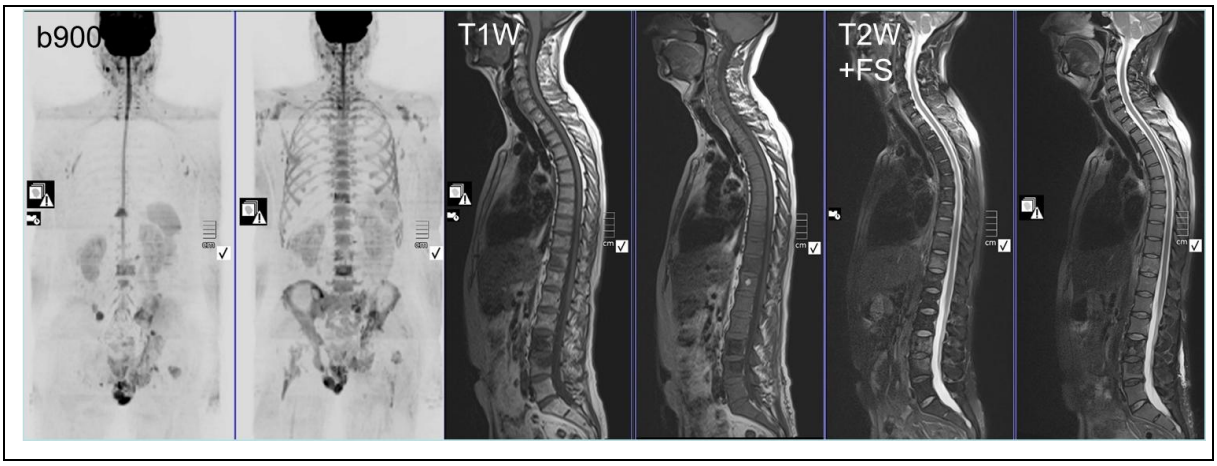
STIR = short tau inversion recovery.

Appendix Table 2 – Discordant assessments of primary, secondary, and tertiary responses		
Primary pattern (first no.)	RAC 1–5	Dominant (≥50%) pattern of response within region
Secondary pattern (second no.)	RAC 1–5	Second most frequent pattern of response within region
Tertiary pattern (third no.)[a]	RAC 4–5	Evidence of progressing disease (only if not primary or secondary pattern)
RAC = response assessment category. [a] The five potential response assessment categories are defined as: (1) highly likely to be responding, (2) likely to be responding, (3) stable, (4) likely to be progressing, and (5) highly likely to be progressing. The maximum % of tertiary pattern 4 or 5 is 20%		

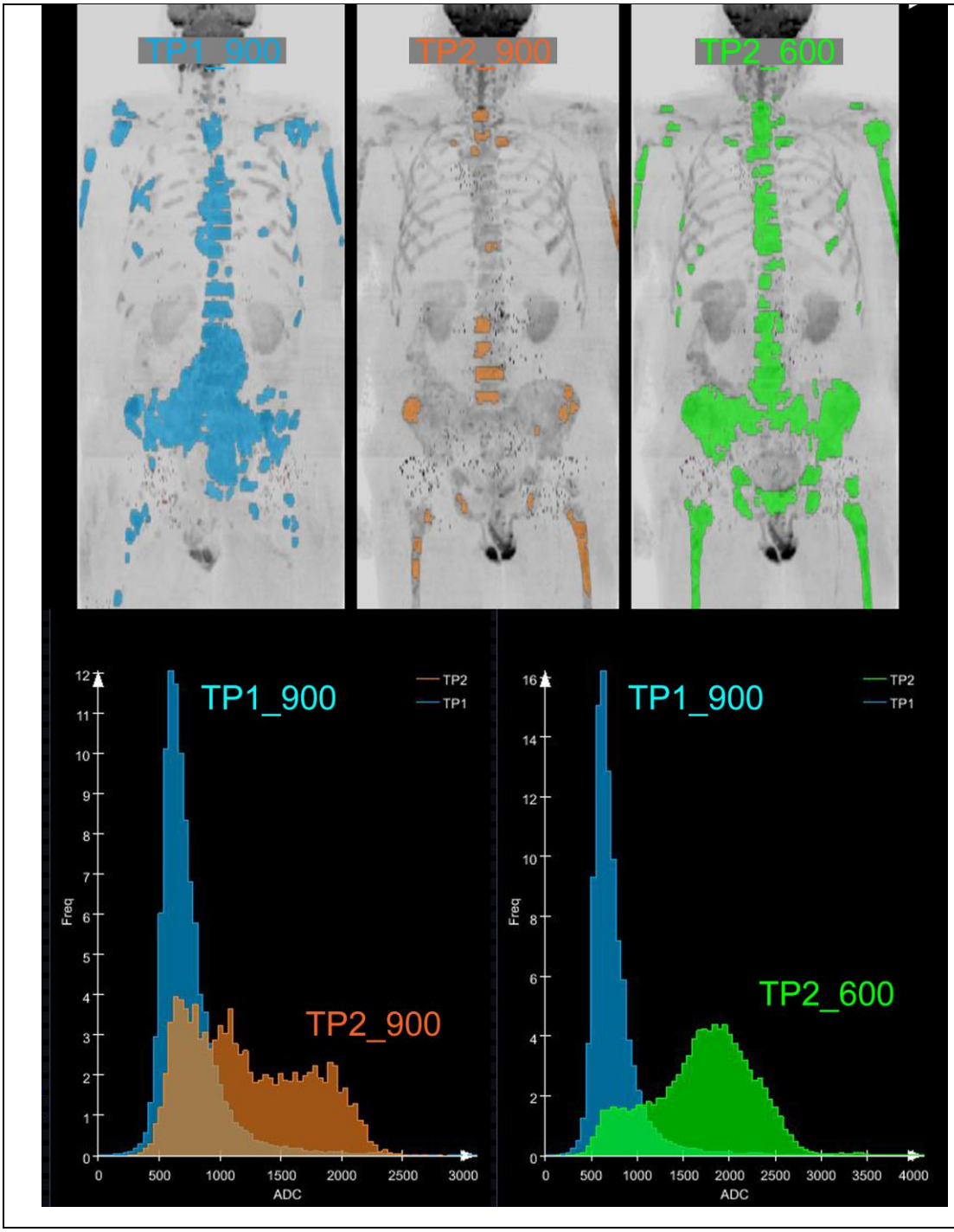
Appendix Fig. 1 – Hardware and software upgrades change image quality. A 51-yr-old woman with metastatic breast cancer in clinical remission, on maintenance trastuzumab emtansine treatment. Three examinations have been obtained on the dates indicated. The dates and sequence parameters have been left on the images for reader review. Images are inverted maximum intensity projection images of b900 diffusion weighted images using inversion recovery fat-saturation sequences. Examination 1 and 2 were undertaken at 1.5T before and after a hardware and software upgrade. Examination 3 was undertaken a few days after Examination 2 at 3T on the same manufacturer's equipment. Note the dramatic difference with the 1.5T upgrade (higher signal-to-noise ratio generally and the improved visibility of anterior structures [anterior ribs and groin nodes]). The examination undertaken at 3T a few days after Examination 2, shows decreased visibility of the normal bone marrow which is due to the effects of increased susceptibility of trabecular bone encountered at 3T.



Appendix Fig. 2 – Bone marrow growth factors obscuring the presence of metastases. A 53-yr-old man with prostate-specific antigen (PSA) oligosecretory metastatic castration-resistant prostate cancer (PSA 6.1 ng/ml), failed abiraterone therapy. Pre- and post-four cycles of docetaxel chemotherapy with granulocyte colony stimulating factor support for all four cycles. PSA is 8.3 ng/ml after treatment. Left image pair: b900 maximum intensity projection projections (inverted scale) shows background bone marrow hyperplasia in the axial skeleton after treatment. Individual deposits are no longer assessable for response. Middle image pair: T1 weighted (W) spine showing the replacement of bone marrow fat consistent with bone marrow hyperplasia. Individual deposits are difficult to assess for response. New lesions cannot be identified with confidence. Right image pair: T2W + fat suppression (FS) showing a mild increase in background signal consistent with bone marrow hyperplasia.



Appendix Fig. 3 – Monitoring therapy response–effects of b-value choice on tumour volume segmentation and apparent diffusion coefficient values. Clinical details: 65-yr-old man with metastatic castrate-resistant prostate cancer being treated with docetaxel chemotherapy for bone and nodal disease. Examinations were obtained at baseline (TP1: prostate-specific antigen 93.1 ng/ml) and after four cycles (TP2: prostate-specific antigen 9.9 ng/ml) on a 1.5T scanner. Segmented whole body maximum intensity projection images at the two time points (TP1 and TP2) with identically applied signal intensity segmentation thresholds. Corresponding apparent diffusion coefficient (ADC) histograms. The x-axis is the ADC value ($\mu m^2/s$) and y-axis is the relative frequency. TP1_900 and TP2_900 are the b900 body maximum intensity projection images with identically applied segmentation thresholds (blue and orange). The histograms of the corresponding outlined tumour volumes are labelled. Note how the application of the threshold to TP1 is successful at outlining bone and pelvic and retroperitoneal nodes. The pretreatment (TP1_900) unimodal histogram (blue) has a median ADC of 739 $\mu m^2/s$ (5th to 95th centile range of 472–1209). The same threshold applied to TP2_900 image results in under-sampling of the bone marrow, with a post-treatment bimodal histogram, median ADC of 1127 $\mu m^2/s$ (5th to 95th centile range of 694–2070). However, the same segmentation threshold applied to b600 body maximum intensity projection images at TP2 (TP2_600) (green) is more successful at sampling the bone marrow, with a resulting post-treatment bimodal histogram, with a median ADC of 1787 $\mu m^2/s$ (5th to 95th range of 687–2502). Note the improved appreciation of tumour response in the histogram shape and summary metrics. This example emphasises the need to carefully choose b-value images that allow adequately sampling of the bone marrow in response assessment settings. Image analysis performed using prototype software (Syngo.via Frontier MR Total Tumor Load; Siemens Healthcare, Erlangen, Germany).


TP1_900
TP2_900
TP2_600
TP1_900
TP2
TP1
Freq
TP2_900
ADC
TP1_900
TP2
TP1
Freq
TP2_600
ADC

Appendix Fig. 4 – Illustrated interpretations of skeletal regional bone assessment form. Example of completed METastasis Reporting and Data System for Prostate Cancer regional follow-up template form 1 (Appendix 5.1) showing the documentation of concordant (eg, pelvis) and discordant (eg, ribs and shoulder girdle) responses. The *comments* are for illustrative purposes and are not required when completing the form.

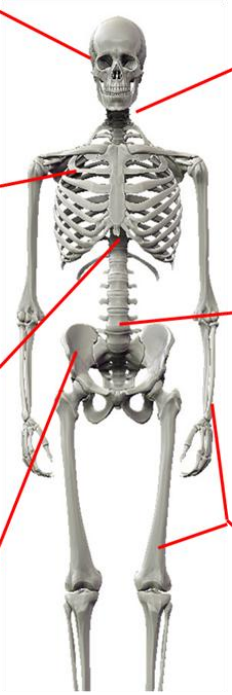
Pt identifier: ..... Scan date: ..../..../.... Scan number: ..... Reporter: .....

SKULL			
Involved	Yes		
Response assessment values	Primary (1-5): 5	Secondary (1-5): -	Tertiary (4-5): -
Comments	Single lesion in skull showing obvious progression		

RIBS AND SHOULDER GIRDLE			
Involved	Yes		
Response assessment values	Primary (1-5): 1	Secondary (1-5): 2	Tertiary (4-5): 5
Comments	Multiple lesions. Majority show clear response. Next most common pattern is likely response. Single lesion shows clear progression (hence tertiary RAC5).		

THORACIC SPINE			
Involved	Yes		
Response assessment values	Primary (1-5): 3	Secondary (1-5): 1	Tertiary (4-5): -
Comments	Multiple lesions. Majority show no change. Next most common pattern is clear response. No tertiary pattern.		

PELVIS			
Involved	Yes		
Response assessment values	Primary (1-5): 1	Secondary (1-5): 1	Tertiary (4-5):
Comments	Multiple lesions all showing signs of clear response. No secondary pattern.		



CERVICAL SPINE			
Involved	Yes		
Response assessment values	Primary (1-5): 1	Secondary (1-5): -	Tertiary (4-5): -
Comments	Single lesion shows complete response.		

LUMBO-SACRAL SPINE			
Involved	Yes		
Response assessment values	Primary (1-5): 2	Secondary (1-5): 1	Tertiary (4-5): -
Comments	Multiple lesions. Dominant change is likely response, but some showing clear response.		

EXTREMITIES (Upper and Lower Limbs)			
Involved	No		
Response assessment values	Primary (1-5):	Secondary (1-5):	Tertiary (4-5):
Comments			