International Journal of Radiation Oncology biology • physics

www.redjournal.org

Critical Review

A Systematic Review of the Clinical Implementation of Pelvic Magnetic Resonance Imaging–Only Planning for External Beam Radiation Therapy

David Bird, MSc,*,† Ann M. Henry, MD,*,† David Sebag-Montefiore, FRCR,*,† David L. Buckley, PhD,‡ Bashar Al-Qaisieh, PhD,* and Richard Speight, PhD*

*Leeds Cancer Centre, Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom; and †Radiotherapy Research Group and ‡Biomedical Imaging, University of Leeds, Leeds, United Kingdom

Received Apr 10, 2019, and in revised form May 22, 2019. Accepted for publication Jun 21, 2019.

The use of magnetic resonance (MR) imaging scans alone for radiation therapy treatment planning (MR-only planning) has been highlighted as one method of improving patient outcomes. Recent technologic advances have meant that introducing MR-only planning to the clinic is becoming a reality, with several specialist radiation therapy clinics using this technique for treatment. As such, substantial efforts are being made to introduce this technique into wide-spread clinical implementation. A systematic review of publications investigating the clinical implementation of pelvic MR-only radiation therapy treatment planning was undertaken following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. The Medline, Embase, Scopus, Science Direct, Cumulative Index to Nursing and Allied Health Literature, and Web of Science databases were searched (timespan: all years to January 2, 2019). Twenty-six articles met the inclusion criteria. The studies were grouped into the following categories: (1) MR acquisition and synthetic computed tomography generation verification, (2) MR distortion quantification and phantom development, (3) clinical validation of patient treatment positioning in an MR-only workflow, and (4) MR-only commissioning processes. Key conclusions from this review are (1) MR-only planning has been implemented clinically for prostate cancer treatments; (2) a substantial amount of work remains to translate MR-only planning into widespread clinical implementation for all pelvic sites; (3) MR scanner distortions are no longer a barrier to MR-only planning, but they must be managed appropriately; (4) MR-only–based patient positioning verification shows promise, but limited evidence is reported in the literature and further investigation is required; and (5) a number of MR-only commissioning processes have been reported, which can aid centers as they undertake local commissioning; however, this needs to be formalized in guidance from national bodies.

Note—An online CME test for this article can be taken at https://academy.astro.org.

Corresponding author: David Bird, MSc; E-mail: David.Bird3@nhs.net

This study is funded by the National Institute for Health Research (NIHR) as part of an HEE/NIHR ICA Programme Clinical Doctoral Research Fellowship (ICA-CDRF-2017-03-005). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Disclosures: No conflicts of interest to declare. D.B. is funded by a National Institute for Health Research Clinical Doctoral Research Fellowship for this research. R.S. is supported by a Cancer Research UK Centres Network Accelerator Award Grant (A21993) to the ART-NET consortium.

Supplementary material for this article can be found at https://doi.org/10.1016/j.ijrobp.2019.06.2530.

https://doi.org/10.1016/j.ijrobp.2019.06.2530

Introduction

One of greatest challenges remaining in radiation therapy is improving the accuracy of treatment volume delineations.[1] Further reducing the uncertainty in delineation could lead to improved patient outcomes by reducing treatment volumes, allowing a reduction in treatment-related toxicities,[2,3] or reducing the risk of geographic misses, thereby improving local control and potentially overall survival rates. The use of magnetic resonance imaging (MRI) scans alone for radiation therapy treatment planning (magnetic resonance [MR]-only planning) has been highlighted as one method of potentially improving target volume delineation accuracy[4-6]; this is due to the improved soft tissue contrast of MRI compared with computed tomography (CT) and the potential to use the other benefits of MRI, such as functional imaging.[5]

Recent technological advances have meant that introducing MR-only planning to the clinic is becoming a reality.[5,7] Hardware and software developments have improved the geometric distortion inherent within MR images to levels that are acceptable for radiation therapy treatment planning,[8] and substantial progress has been made in acquiring electron density information from MRI data alone through synthetic CT generation methods.[4,9] The field of synthetic CT generation has been reviewed,[4,9] and commercial solutions are available, including several prostate solutions and, recently released, a solution for the whole pelvis.[10-12] Consequently, MR-only treatments are being conducted by specialist radiation therapy clinics and over time are likely to move to more widespread clinical implementation.

This systematic review assesses the literature surrounding the clinical implementation of pelvic MR-only radiation therapy treatment planning with the aim of detailing and discussing the breadth of work that has been undertaken. This review considers only work that has been published in relation to MR-only planning for pelvic external beam radiation therapy.

Methods and materials

A systematic review of publications investigating the clinical implementation of pelvic MR-only external beam radiation therapy treatment planning was performed using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.[13] The Medline, Embase, Scopus, Science Direct, Cumulative Index to Nursing and Allied Health Literature, and Web of Science databases were searched with a time span of all years to January 2, 2019 (for Medline and Embase, this corresponded to "Week 3 December 2018" and "Week 52 2018," respectively) using the search protocols in Appendix E1 (available online at https://doi.org/10.1016/j.ijrobp.2019.06.2530).

Articles were included that referred to "MR-only" or "synthetic-CT" and "radiotherapy" or synonyms of these terms in their title or abstract. These deliberately broad search criteria were used to minimize the risk of relevant studies not being identified. The search results for each database were combined, and duplicates were removed. The remaining results were screened from their titles for eligibility. Primary screening included only search results that were related to the use of MRI in radiation therapy for cancer treatment. Secondary screening included only articles related to the clinical implementation of MR-only external beam radiation therapy treatment planning for pelvic cancer sites. Articles focusing on the general use of MRI in radiation therapy, MRI in brachytherapy, synthetic CT model generation techniques, target volume delineation using MRI, MR image registration, positron emission tomography/MRI in radiation therapy, MRI safety, MR-only contouring, and MR-only fiducial marker identification were excluded. Articles regarding synthetic CT model generation techniques were specifically excluded because they were appraised recently in the literature within 2 review articles[4,9] and because this review is focused on the clinical implementation, rather than the technique development, aspect of MR-only planning. Conference proceedings were excluded because of their large number and variable information provision, which made their inclusion unbeneficial. A backward citation search of the remaining eligible studies was undertaken. The included studies were categorized according to their focus. For each category, key findings from each study were included in a data table.

Results

The database search results can be seen in Figure 1. The combined database search resulted in 2024 records, with 1066 records remaining after duplicate removal. After primary screening, 535 records remained. After secondary screening, 71 studies remained. After further review, 49 studies were removed (44 conference abstracts and 5 studies that did not meet the eligibility criteria). Twenty-two studies remained, to which the citation search added 4 studies; thus, 26 studies were included in this systematic review. The categories and number of excluded articles can be seen in Table 1.

Included studies were organized into one or more of the following categories for review: (1) MR acquisition and synthetic CT generation verification, (2) MR distortion quantification and phantom development, (3) clinical validation of patient treatment positioning in an MR-only workflow, and (4) MR-only commissioning processes. These 4 categories will be discussed in more detail.

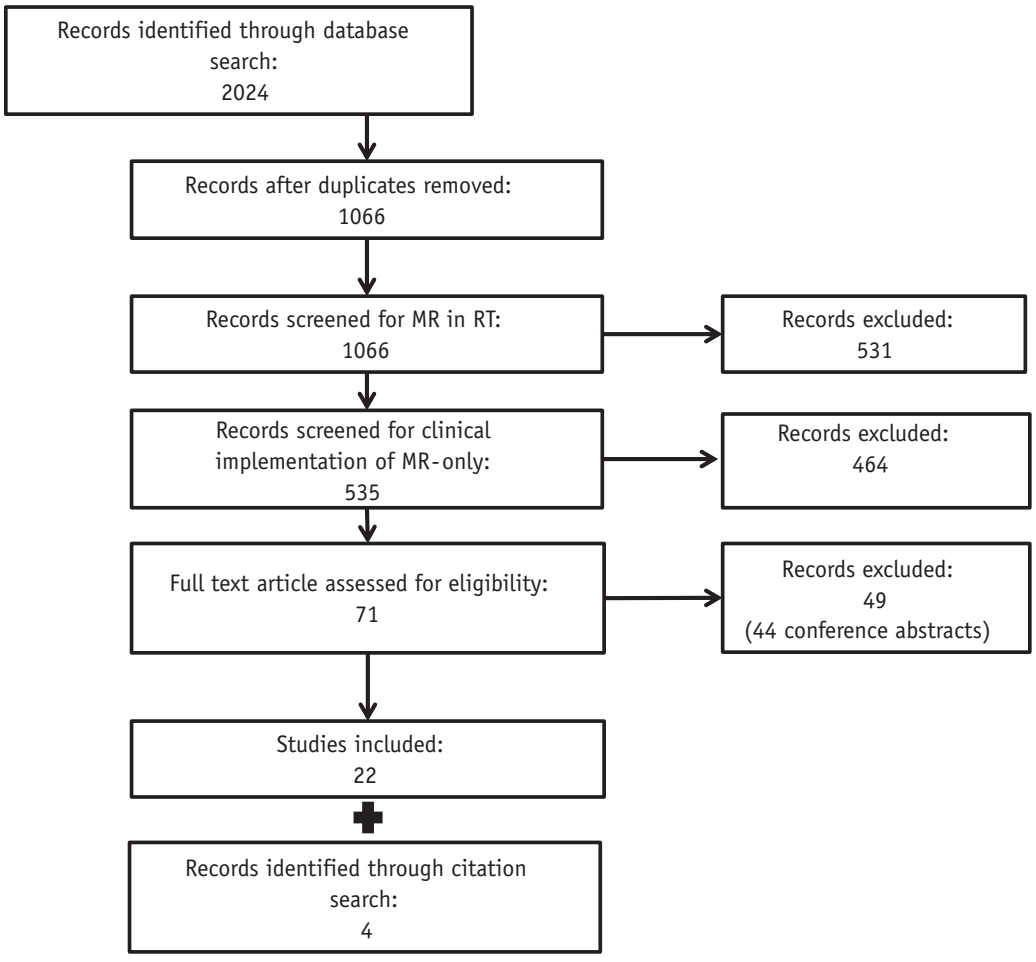

Fig. 1. Flowchart of the systematic review process, including the number of studies included in this review.

MR acquisition and synthetic CT generation verification

The systematic review identified 9 studies investigating MR acquisition or synthetic CT generation verification.[6,14-21] A key summary of study results can be found in Table 2. All studies reported results relating to prostate cancer treatment planning. Studies reporting synthetic CT dosimetric accuracy findings were included when they were validating previously reported synthetic CT generation models as part of clinical implementation, rather than as part of the development of a synthetic CT model.

Tenhunen et al,[14] Kerkmeijer et al,[15] Christiansen et al,[16] and Tyagi et al[17] reported treating patients using an MR-only pathway, with the number of patients treated ranging from 125 to 1.[14,16] Tenhunen et al[14] also reported the MR-only patient cohort's initial clinical response in terms of early response PSA and acute toxicity follow-up.

Persson et al,[6] Tenhunen et al,[14] and Tyagi et al[17] reported their experiences of prospectively acquiring MR data for MR-only treatment planning in terms of their MR scan success rates and the issues that prevented successful scanning. In the case of Persson et al,[6] case, this was from a multicenter research-only study for commissioning purposes, whereas other experiences[14,17] resulted from treating their first MR-only patients. Wyatt et al[20] reported MR scan success rates from retrospectively assessed MR data, whereas Christiansen et al[16] reported synthetic CT generation success rate but did not discuss issues regarding MR acquisition. In describing their clinical workflow for MR-only planning, Tyagi et al[17] also reported a time savings when using an MR-only versus CT-MRI—based workflow.

Previous studies[6,16,18-21] reported validating the dosimetric accuracy of their respectively chosen synthetic CT solution in a clinical environment, as required for commissioning MR-only planning.

MR distortion quantification and phantom development

The systematic review identified 13 studies investigating MR distortion quantification methods or phantom development.[18,19,22-32] A key summary of study results can be found in Table 3.

Table 1 Categories and number of articles excluded from this review after primary, secondary, and tertiary screening

Reasons for exclusion	No. of articles
Primary screening	
Not related to cancer treatment	65
Not related to radiation therapy cancer treatment	301
Not related to the use of MRI in radiation therapy cancer treatment	165
Total	531
Secondary screening	
General use of MRI	115
Other site synthetic CT generation technique	98
Brain synthetic CT generation technique	65
Brachytherapy, Gamma knife	48
Prostate synthetic CT generation technique	43
MR delineation	33
MR image registration	12
Proton or ion synthetic CT generation technique	12
Other	10
MR-only contouring	7
MR-only fiducial marker identification	7
PET-MRI	7
UTE synthetic CT generation techniques	5
MR safety	2
Total	464
Tertiary screening	
Conference abstracts	44
MR-only review articles	3
Synthetic CT model development studies	2
Total	49

Abbreviations: CT = computed tomography; MR = magnetic resonance; MRI = magnetic resonance imaging; PET = positron emission tomography; UTE = ultrashort echo time.

Several authors[22-26] developed phantoms for use in measuring geometric distortions or end-to-end testing the MR-only pathway. Huang et al,[23] Price et al,[22] and Walker et al[24] developed large field of view (FoV) phantoms for assessing system (Bo) distortions and characterized their respective MR scanner distortions. Price et al[22] and Huang et al[23] assessed the setup reproducibility of their phantoms using CT scan testing methods, and Walker et al[24] assessed the effects of a continuous moving-table acquisition method on measured distortions with 0, 1.1, and 2 mm/s table speeds.

Sun et al[25] and Cunningham et al[26] developed anthropomorphic pelvic-shaped phantoms for measuring system and patient-induced susceptibility distortions or end-to-end testing of the MR-only pathway. Both phantom designs were based on prostate patient anatomic sizes. The phantom of Sun et al[25] had 2 designs for end-to-end testing or geometric distortion assessment. Cunningham et al[26] designed the phantom so that it could simulate patient bladder and rectal filling for end-to-end testing, including dosimetric verification of treatment plans.

Some authors[18,19,27-29] investigated the effects of MR scanner distortions on patient treatments by applying measured or simulated distortions to patient treatment plans. Kemppainen et al[19] and Gustafsson et al[27] measured system-induced geometric distortions using large FoV phantoms for 15 and 10 patients, respectively. Tyagi et al,[18] Kemppainen et al,[19] and Glide-Hurst et al[28] measured patient-induced susceptibility geometric distortions for 20, 4, and 9 patients, respectively, with Glide-Hurst et al[28] assessing distortions for different patient bladder-filling states and scanner magnet strengths. Adjeiwaah et al[29] assessed the effects of MRI scanner-measured system and simulated patient-induced susceptibility distortions for 17 patients.

Wyatt et al[30] evaluated the repeatability and setup sensitivity of the commercially available GRADE (Spectronics Medical AB, Helsingborg, Sweden), large-FoV distortion phantom. The distortion measurement repeatability was assessed for interscanning and intrascanning session variability. The setup sensitivity of the phantom was investigated by deliberately scanning the phantom with a 1-mm offset and 1 degree of rotation and assessing distortion variations.

Torfeh et al[31] and Price et al[32] characterized their MR scanner system and gradient nonlinearity distortions respectively over large FoVs as required for MR-only planning. Torfeh et al[31] assessed the effects of manufacturer 2-dimensional and 3-dimensional distortion correction algorithms on clinically used radiation therapy sequences. Price et al[32] characterized and minimized inherent 2-dimensional and 3-dimensional large-FoV gradient nonlinearity distortions using postprocessing techniques.

Clinical validation of patient treatment positioning in an MR-only workflow

The systematic review identified 3 articles investigating the clinical validation of patient treatment positioning verification.[18,33,34] A key summary of study results can be found in Table 4. These studies have been included because they report patient treatment positioning verification results for previously reported MR-only treatment synthetic CT models as part of clinical implementation.

Tyagi et al,[18] Kemppainen et al,[33] and Korhonen et al[34] evaluated the accuracy of synthetic CTs as digitally reconstructed radiograph (DRR) reference images for treatment positional verification using orthogonal planar images,[18,33,34] cone beam CT (CBCT), or both.[18,34]

Tyagi et al[18] and Kemppainen et al[33] investigated the Philips MRCAT synthetic CT solution. The DRR analysis by Kemppainen et al[33] included interobserver and intraobserver variability, separating the variability into systematic and random error contributions and comparing the total geometric accuracy to a reference of 2 mm error from CT to MR registration. The CBCT analysis by Tyagi et al[18] was based

Table 2 Summary of the key results from the MR acquisition and synthetic CT generation verification studies*

Author	Year	sCT technique	No. patients in study	MR acquisitions	MR scanner and magnet strength	No. of patients treated with MR only	sCT success rate (%)	Other key information
Christiansen et al[16]	2017	Philips - MRCAT	30	T1 mDIXON	Philips Ingenia, 1.5 T	1	97 (29/30)	1. MR synthetic CT generation failed in 1 case, reason unknown 2. Dosimetric accuracy for gamma analysis of 2%/2 mm - median 100% in all structures 3. Rectal gas found to be main contributor to dosimetric errors
Kemppainen et al[19]	2017	Philips - MRCAT	5	T1 mDIXON	Philips Ingenia, 1.5 T	X	X	Only prostate patient data from study included 1. Mean dosimetric accuracy (prostate patients) for 2%/2 mm and 1%/1 mm gamma analysis of 100% and 99.2%, respectively, within PTV 2. Mean relative dose difference of 0.7% in PTV and <1.2% in OARs
Maspero et al[21]	2017	Philips - MRCAT	14	T1 mDIXON	Philips Ingenia, 3 T	X	X	Mean relative dose difference between CT and sCT found to be 0.3% within the CTV and 0.04% within the whole body
Persson et al[6]	2017	Spectronics Medical - MRIPlanner	170	T2 SPACE	GE Discovery (3 T), GE Signa (3 T), Siemens Area (1.5 T), Siemens Skyra (3 T)	X	85 (145/170)	1. Patient MR acquisition issues (no. of patients): distortion correction turned off (12), whole body not included in FoV (4), insufficient superior-inferior coverage (2), hip prosthesis patients (2), extreme rectum change between CT and MR (1) 2. Mean dosimetric deviations of less than 0.3% for all targets and organs 3. Multicentered (4 centers) study found insignificant differences found between range of treatment techniques, planning systems, prescribed doses, calculation models and target volumes
Tyagi et al[17]	2017	Philips - MRCAT	48	T1 mDIXON	Philips Ingenia, 3 T	42	87.5 (42/48)	1. Patient MR acquisition issues (no. of patients): hip prosthesis patients (4), large patient exceeded MRCAT size limitations (2) 2. Dedicated software used ed for contouring workflow. MR sequence blurring affected 2-dimensional DRR fiducial marker identification in 2 patient cases 3. MRCAT failure modes: (i) presence of hip prosthesis, (ii) significant bone disease in pelvis, (iii) significant discrepancies from the bone model boundary conditions, and (iv) patient size exceeds 50 cm left-right or 30 cm anteroposterior 4. Time saving of ~15 minutes using MR-only simulation compared with CT-MR simulation, further 15-minute savings estimated in the future if logistic challenges resolved
Tyagi et al[18]	2017	Philips - MRCAT	25	T1 mDIXON	Philips Ingenia, 3 T	X	X	Mean relative dose difference between sCT and CT <0.5%
Wyatt et al[20]	2017	Dowling et al[40]	21	T2 SPACE	Philips Magneto Espree, 1.5 T	X	54 (21/37)	Retrospective data collection 1. Patient data set exclusions (no. of patients): required patient body outside of MR FoV (13), hip prostheses (2), gross patient motion (1) 2. Dosimetric accuracy for 2%/2 mm gamma analysis: mean 98.9%, minimum 97.6%, and maximum 99.5% in all structures

(*continued on next page*)

Table 2 (*continued*)

Author	Year	sCT technique	No. patients in study	MR acquisitions	MR scanner and magnet strength	No. of patients treated with MR only	sCT success rate (%)	Other key information
Kerkmeijer et al[15]	2018	Philips - MRCAT	Not known	T1 mDIXON	Philips (unknown)	Yes, number unknown	Not known	1. Inclusion criteria: fiducial markers present in prostate 2. Exclusion criteria: hip prostheses and contraindications for MRI
Tenhunen et al[14]	2018	Korhonen et al[39]	250	T1 mDIXON	GE Optima, 1.5 T	125	92 (184/200)	1. Patient MR acquisition issues (no. of patients): gold markers not identifiable (8), hip prosthesis related distortions (5), obesity (2), motion (1) 2. CT vs MR-only patient treatment outcomes: PSA and acute toxicities results showed no significant differences between pathways 3. Noted lack of support of MR-only workflow from technical software, including planning systems

Abbreviations: CT = computed tomography; DRR = digitally reconstructed radiograph; MR = magnetic resonance; MRI = magnetic resonance imaging; PTV = planning target volume; sCT = synthetic computed tomography.

* "No. of patients in study" refers to the total number of patients recruited for MR-only investigations, "MR-only treated patients" refers to the number of patients planned and treated using the MR-only technique, and "sCT success rate" refers to the percentage of patients for whom a useable sCT was generated.

on fiducial marker 3-dimensional CBCT scans, in which 5 CBCT scans were included for registration per patient.

The DRR analysis by Korhonen[34] included interobserver variability and investigated the use of both the synthetic CT and MR images as reference images for CBCT registration. CBCT registrations were undertaken using ELEKTA x-ray volume imaging (XVI) software (Elekta, Stockholm, Sweden) for 5 patients, with 10 CBCT scans for each patient (50 CBCT registrations per reference modality).

MR-only commissioning processes

The systematic review identified 6 articles investigating MR-only commissioning processes.[15,21,35-38] A key summary of study results can be found in Table 5.

Kerkmeijer et al,[15] Kapanen et al,[35] and Kim et al[36] reported experiences related to commissioning an MR-only pathway. Kerkmeijer et al[15] and Kapanen et al[35] used their experiences of commissioning an MR-only pathway and an MR simulator, respectively, to present recommendations for clinically commissioning an MR-only pathway, including proposing quality-assurance testing and associated levels of acceptability with individual pathway components. Kim et al[36] used a failure mode and effects analysis methodology to systematically assess the risks—and their frequency, severity, and detectability—of an MR-only planning pathway compared with CT-MR based pathway. This analysis included mapping the respective elements required for an MR-only pathway, their risks, and associated mitigation strategies.

Maspero et al[21] and Korsholm et al[37] reported synthetic CT accuracy assessment methodologies. Maspero et al[21] quantified the confounding factors in MR-only dose calculation accuracy assessments for patients with prostate cancer, including interscan differences (setup and positioning differences, MR-related geometric inaccuracy, and registration errors) and synthetic CT generation and electron density conversion errors. Korsholm et al[37] developed a statistical approach to evaluating the significance of errors introduced by MR-only planning compared with CT-based planning, with the criterion that 95% of patients should have an uncertainty in dose calculation within 2% of the CT dose for relevant structures.

Palmer et al[38] developed and validated a quality assurance procedure for assessing synthetic CT clinical feasibility using kV-CBCT, where CBCT scans were used to recalculate the synthetic CT treatment plan as a check of its dose calculation accuracy.

Discussion

A wide range of findings are reported in this systematic review. Several key findings are seen in the literature, and these are highlighted here before being discussed in more detail later. These findings are (1) MR-only planning has been clinically implemented for prostate cancer treatments; (2) a substantial amount of work remains to translate MR-only planning into widespread clinical implementation for all pelvic sites; (3) MR scanner distortions are no longer a barrier to MR-only planning, but they must be managed appropriately; (4) MR-only based patient positioning verification shows promise, but limited evidence is reported in the literature and further investigation is required; and (5) a number of MR-only commissioning processes have been

reported, which can aid centers as they undertake local commissioning, but this needs to be formalized in guidance from national bodies.

As highlighted, in 4 studies patients with prostate cancer were treated using an MR-only planning solution, showing that clinical implementation is achievable.[14-17] It is interesting to note that all commissioning work identified in this review was also focused on prostate treatments. This is a natural starting point for pelvic MR-only planning because other pelvic sites (rectum, bladder, anus, gynecological) have a number of additional challenges associated with them, including differences in male and female anatomy, significantly larger treatment volumes, and non−fiducial-marker-based 3-dimensional imaging requirements, which makes their implementation more complex. It is important to note that the majority of work discussed is translatable to other cancer sites; however, it is clear that a significant amount of work remains to widen the implementation of MR-only planning to all pelvic cancer treatments.

This review identified 3 key areas that were investigated for clinical implementation purposes: MR acquisition and synthetic CT generation verification, MR distortion quantification and phantom development, and clinical validation of patient treatment positioning in an MR-only workflow. In each, no major barriers to implementation were found, and a number of publications reported commissioning methodologies that will benefit the wider community by providing guidance for local centers to use within their own MR-only clinical commissioning.

The first step to implementing an MR-only pathway is ensuring that sufficient and suitable MR data acquisition is achieved. A high success rate of acquiring MRI that is usable for synthetic CT generation is key to the widespread implementation of MR-only techniques because this will limit the need for additional CT scans. Persson et al,[6] Tenhunen et al,[14] and Tyagi et al[17] all described their success rates in prospective studies and categorized the identified issues related to scanning. The similar scan success rates (85%-92%) suggest that this is an achievable percentage in any center, particularly considering the multicenter study by Persson et al.[6] The differences in success rate can be explained by the variations in study design. The exclusion criteria of Persson et al[6] included patients with hip prostheses and operator error as valid reasons for an unsuccessful MRI scan, whereas Tenhunen et al[14] and Tyagi et al[17] had no exclusion criteria. Wyatt et al[20] also analyzed their successful scanning rate (54%); however, their rate was severely affected by a lack of scanning guidance for MR operators caused by the retrospective design of the study. However, the study does still provide useful information regarding common issues with MR acquisition in this context. Centers should ensure that training from experienced personnel is provided for MR scan operators and consider methods to identify errors at the point of acquisition to ensure MR scan success. Tenhunen et al[14] and Tyagi et al[17] identified several patient and hardware- and software-related issues that also prevented successful MR-only planning and therefore required a percentage of patients to revert to a CT-MRI−based pathway. Although further development of MR-only solutions could lead to a reduction in patients requiring an additional CT scan, provisions should still be made for CT-based planning to occur. In addition, these studies do not discuss patients who have contraindications to MRI and therefore will always require a CT-only pathway.

Christiansen et al,[16] did not report MR acquisition success, but they described their synthetic CT generation failure rate, finding that 3% (1 of 30) of synthetic CT examinations failed to generate using the Philips MRCAT solution. This was considered to be due to the software's "sanity" check ability to prevent obviously erroneous synthetic CT generation, although the exact cause was not established. This finding highlights that synthetic CT generation methods require input data to follow clearly defined criteria to be successful, and it is a beneficial feature that the Philips MRCAT safeguards against inappropriate data, defined as including large patient sizes, large disease sites (300 mm or greater scan lengths), and hip prostheses.[17,19] It is of note that Tyagi et al[17] did not have any similar issues. This result could have been due to a systematic difference in pathway, such as Tyagi's use of a specific mold for each patient to achieve a more robust patient position, or a non-systematic, patient-specific, issue. This is another example of the variety of errors associated with an MR-only planning pathway that require careful assessment.

For acquired MR data to be clinical usable, their dosimetric accuracy needs to be quantified robustly as within acceptable limits. Dosimetric accuracy of prostate synthetic CT solutions was investigated by the majority of studies and considered to be clinically acceptable in all cases.[6,16,18-21] The similar results of these studies, despite significant differences in study design including various synthetic CT generation methods, shows that the dosimetric accuracy of synthetic CT techniques is broadly reproducible across a wide range of clinical systems and techniques, including multiple commercial options available for prostates.[10,11] The presented studies also provide a suitable blueprint for centers wishing to begin clinical implementation of MR-only planning themselves regarding dosimetric accuracy assessment. These studies progress by first assessing dosimetric accuracy through research-only studies, followed by end-to end pathway testing and eventual implementation only when the local results provided sufficient confidence that the MR-only planning was sufficiently dosimetrically accurate to be used without CT for assurance.

In addition to dosimetric accuracy, another key criterion of useable MR data for MR-only treatment planning is that it be geometrically accurate. To be sufficiently confident of this for clinical implementation requires robust quality assurance techniques, phantoms, and the characterization of the MRI distortions. The reported studies focused on designing suitable phantoms for measuring geometric distortions or end-to-end testing the MR-only pathway,[22-26]

Table 3 Summary of the key results from the MR distortion quantification and phantom development studies

Author	Year	Phantom and software	Phantom Shape (cm × cm × cm)	MR scanner and magnet strength	Other key information
Price et al[32]	2015	Philips temporal GNL phantom; in-house 3D distortion phantom; in-house software	2D: 36 × 43 × 2 3D: 46.5 × 35 × 16.8	Philips Panorama, 1 T	1. Gradient nonlinearity distortions found to be stable over 6-month period 2. Vendor 3D distortion corrections maintained <1 mm distortion up to 9.5 cm from isocenter 3. Postprocessing corrected distortions <1 mm for large FoV up to 25 cm from isocenter 4. Significant inherent gradient nonlinearity distortions may be a specific feature of open-bore MR scanners, rather than cylindrical scanners, due to shorter gradient coils
Sun et al[25]	2015	Self-developed pelvic-shape phantoms and software	25 × 40 × 26	Siemens Skyra, 3 T	1. Phantom internal details: spherical and cylindrical inserts representing prostate, rectum bladder, and femoral heads based on average of 39 prostate patients or 11 plastic grid sheets 2. Maximum distortion across phantom with 3D correction found to be 1.7 mm (75% quartile, 0.54 mm). 3. Phantom end to end testing found mean dose difference of 1.1 cGy between CT and MR
Walker et al[24]	2015	Self-developed large FoV phantom and software	Max: 50 × 50 × 51.3	Siemens Skyra, 3 T	1. Maximum 3D distortion correction distortion was 4.08 mm for a 2-mm SE sequence. 2. Within 152 mm of isocenter for 2 mm SE with 3D distortion correction, distortion ≤ 2 mm 3. For the continuous moving-table mode, 1.1 mm/s was found to have the least distortion with a maximum of 4.4 mm and a distance of 140 mm within which the distortion was less than 2 mm.
Huang 7et al[23]	2016	Self-developed large FoV phantom and software	46.5 × 35 × 16.8	Siemens Skyra, 3 T	Mean Bo distortion <1 mm found within a radius of 15 cm from the isocenter
Torfeh et al[31]	2016	GE large-FoV phantom and in-house software	50 × 50 × 50	GE MR-Sim, 1.5 T	1. In-house software validated with a mean distortion error of 0.15 mm 2. Mean Bo distortion both in-plane and through plane found to be <2 mm within a radius of 25 cm when manufacturer 2D and 3D distortion applied as recommended 3. Without distortion correction, the size of distortions made use for radiation therapy purposes unachievable
Gustafsson et al[27]	2017	Spectronics large-FoV GRADE phantom and software	50.2 × 40.4 × 53.4	GE Discovery, 3 T	1. Mean and maximum distortions <0.5 mm and <12.6 mm, respectively 2. Maximum distortions: 0.43 mm at <100 mm, 0.82 mm at 100-150 mm, 1.85 mm at 150-200 mm, and 7.9 at 200-250 mm, increasing with radial distance from isocenter 3. Structure deformation was minimal with mean magnitude 0.01 mm for internal structures and <0.33 mm for the full-body contour; mean percentage dose difference was ±0.02%.
Kemppainen et al[19]	2017	Large-FoV phantom and software, unknown origin	Minimum: 37.5 × 37.5 × 45.5	Philips Ingenia, 1.5 T	1. Mean system distortion of <1 mm measured within all PTVs with mean maximum distortion within patient body contours of <2 mm 2. Effects of geometric distortion on dose calculation accuracy found to be <0.2% for all PTVs, with mean patient-induced distortions <1 mm in all cases

(*continued on next page*)

Table 3 (*continued*)

Author	Year	Phantom and software	Phantom Shape (cm × cm × cm)	MR scanner and magnet strength	Other key information
Price et al[22]	2017	Self-developed large-FoV phantom in-house software	Maximum: 55 × 55 × 45	Philips Panorama, 1 T; Philips Ingenia, 1.5 T; Philips Ingenia, 3 T	1. Phantom modular to allow variation in setup for different scanners 2. Setup reproducibility measured to be 0.1, 0, and −0.6 mm respectively in X, Y, and Z directions with negligible rotations 3. Distortion <1 mm within 100 mm radially to isocenter
Tyagi et al[18]	2017	X	X	Philips Ingenia, 3 T	Mean patient-induced susceptibility geometric distortion of −0.07 mm (range, −0.73 to −0.56 mm) and −0.2 mm (range: −0.62 to −0.35 mm) within the outer body and prostate, respectively
Adjeiwaah et al[29]	2018	Spectronics phantom and software	35.1 × 47 × 45.1	GE Signa, 3 T	1. For sequences of bandwidths of 122, 244, and 388 Hz, system distortions were <3.19 mm, <2.52 mm, and <2.08 mm within a radial distance of 25 cm from the isocenter, and the patient-induced distortions were <5.8 mm, <2.9 mm, and <1.5 mm, respectively 2. Dosimetric analysis found that a mean dose difference of <0.5% was found between distorted and undistorted treatment plans 3. Higher bandwidth sequences are recommended to minimize distortion effects
Cunningham et al[26]	2018	Self-developed male pelvic-shape phantom	23 × 38.1 × Unknown	Not applicable	1. External and internal organ shapes based on data from 19 prostate cancer patients 2. Internal structure: pelvic bone anatomy, prostate, urethra, and fillable bladder and rectum 3. Modular changes are possible to accommodate dosimetry inserts or organ changes 4. Phantom able to accurately and reproducibly simulate rectum and bladder filling and to dosimetrically verify treatment plans, with an assessment plan found to have a dose difference of 1.5% between the calculated and measured doses
Glide-Hurst et al[28]	2018	X	X	Philips Panorama, 1 T; Philips Achieva, 1.5 T; Philips Ingenia, 3 T	1. Empty, partially full, and full bladder states investigated over ∼45 minute scanning session 2. Patient-induced susceptibility distortions were small with <2% of prostate and seminal vesicles voxels distorted by >0.5 mm and all-bladder voxels distorted by <1 mm. 3. A significant change in rectal gas seen to increase distortion
Wyatt et al[30]	2018	Spectronics Large FoV GRADE phantom and software	50.2 × 40.4 × 53.4	Siemens Magnetom Espree, 1.5 T; Siemens Prisma, 3 T; GE Signa PET-MR, 3 T	1. Bo distortion measurements for intrascanning and interscanning sessions were repeatable 2. Mean range of measurement for all scanners and sequences less than 1 mm, maximum ranges 2.9 mm and 2.6 mm for 1.5-T and 3-T scanners 3. Phantom found to be relatively sensitive to large set up errors ∼1 mm translation or 1 degree of rotation

Abbreviation: 2D = 2-dimensional; 3D = 3-dimensional; FoV = field of view; MR = magnetic resonance; PTV = planning target volume; SE = spin echo.

the quantification of distortions on patient data,[18,19,27-29] or the reproducibility of distortion measurements[30] and provide information to aid distortion commissioning for an MR-only pathway.

Distortions within 1-T,[22,28,32] 1.5-T,[19,22,28,30,31] and 3-T scanners[18,22-25,27,28,30] from a range of manufacturers, including Siemens,[23-25,30] Philips,[18,19,22,28,32] and GE,[27,31] were measured within a satisfactory range for MR-only

Table 4 Summary of the key results from the clinical validation of patient treatment positioning in MR-only workflow studies

Author	Year	sCT technique	No. of Patients	DRR/ CBCT	2D method	CBCT method	Inter/ Intraobserver
Korhonen et al[34]	2015	Korhonen et al[39]	DRR, 5; CBCT, 5	DRR and CBCT	Manual, bony registration	Automatic (bony and gray) value, 3D and 6D registration	DRR interobserver: 10
Tyagi et al[18]	2017	Philips MR-CAT	DRR, 20; CBCT, 5	DRR and CBCT	Manual, fiducial marker registration	Manual, fiducial marker registration	X
Kemppainen et al[33]	2018	Philips MR-CAT	20	DRR	Manual, bony registration	X	Interobserver, 5; intraobserver, 3

Abbreviations: 2D = 2-dimensional; 3D = 3-dimensional; 6D = 6 degrees of freedom; AP = anterior-posterior; CBCT = cone beam computed tomography; CT = computed tomography; DRR = digitally reconstructed radiograph; LR = left-right; MR = magnetic resonance; sCT = synthetic computed tomography; SI = superior-inferior.

planning, considered to be 2 mm.[8] Unlike the other presented studies, Wyatt et al[30] and Price et al[32] suggested that the majority of distortions measured as part of their studies were larger than clinically acceptable for MR-only planning. However, both authors noted that their MR sequences were not optimized for clinical use because their acquired sequence bandwidths were insufficient to reduce distortion to within acceptable levels, where a minimum suitable bandwidth is considered to be 220 Hz/pixel at 1.5 T and 440 Hz/pixel at 3 T. This is a timely reminder that scanners and clinically used scans require distortion characterization to ensure they are suitable for use.

It is important that distortions be placed in context by evaluating their effect on patient treatments. A number of studies did this by assessing the effects of different distortions when applied to patient treatment plans, with Kemppainen et al[19] and Gustafsson et al[27] assessing system distortions; Tyagi et al,[18] Kemppainen et al,[19] and Glide-Hurst[28] assessing patient-induced susceptibility distortions; and Adjeiwaah et al[29] assessing both system and patient-induced susceptibility distortions. These results broadly showed that these distortions can be considered to have negligible effects on the patient plan, although their assessment and subsequent protocol optimization is vital. This is particularly true when regarding patient-induced distortions, which cannot be corrected systematically because they vary between patients; however, these studies provide confidence that their effects can be quantified or negated for a range of scanners and magnet strengths. Similar investigations should form part of any center's clinical implementation of MR-only planning to allow local distortion effects to be quantified and assessed as clinically significant or not. It is noted these studies were undertaken with low patient numbers (20 or fewer), and in none of the studies were distortions correlated with patient size. Because distortions increase with distance from the isocenter, their effects will increase with patient size. Without information relating to patient sizes, it is not possible to assess whether the true effect of the distortions on larger patients has been quantified. Selection of a large range of patient sizes and quantification of the impact of distortion as a function of patient size potentially would be of more value to a commissioning center than an attempt to establish the "average" patient size.

It is also to the wider radiation therapy community's benefit that self-developed phantom designs be detailed in the literature[22-26] to allow centers to replicate these phantoms. It was noticeable that for pelvic MR-only large-FoV distortion measurements, these studies used only 2 commercially available phantoms—the Spectronics Medical's GRADE[27,30] and GE's large-FoV phantoms[31]—with the remaining studies developing their own phantoms.[22-26] However, it should be noted that there are several other commercially available phantoms that have not been reported here, including the Quasar MRID 3D, CIRS large FoV, Phantomlab MagPhan RT, and Philips MRI distortion phantoms. It is possible that

Table 4 Summary of the key results from the clinical validation of patient treatment positioning in MR-only workflow studies *(continued)*

	Other key information	
CBCT (maximum difference)	Gray value method	sCT vs CT—2 mm (3D), and 1.7 mm, 1.1° (6D) MR vs CT—4 mm (3D), and 3.5 mm, 1.6° (6D)
	Bone method	sCT vs CT—1.6 mm, and 1.3° (6D)
DRR	Heterogeneous sCT vs CT—manual registration errors were highest in the PA direction with mean differences of -0.3 ± 1 m and 0.3 ± 1.7 mm for kV and MV acquired positional images, respectively	
CBCT (mean difference)	sCT vs CT: $<1 \pm 0.79$ mm, $<1 \pm 0.89$ mm, $<0.5 \pm 0.85$ mm for LR, AP, and SI directions, respectively	
Other information	Individual registration differences were observed up to 2 mm in some fractions with larger variations in prostate rotation	
DRR (mean difference)	sCT vs CT: 0.3 mm, 0.3 mm, and 0.6 mm in the lateral, vertical and longitudinal directions, respectively	
DRR (mean difference)	sCT vs CT: −0.5 mm, +0.1 mm and +0.1 mm in the vertical, longitudinal, and lateral directions, respectively	
Other information	1. Repeatability coefficients were 2.1 mm vs 2.6 mm, 1.4 mm vs 2.1 mm, and 1.2 mm vs 1.4 mm in vertical, longitudinal, and lateral directions between CT and sCT, respectively 2. Significant increase in intraobserver variability found for vertical, longitudinal, and lateral directions; however, magnitude was less than 0.5 mm in all directions 3. MRCAT has positive effect on total geometric accuracy compared with 2-mm registration error of CT-MR pathway	

the use of these phantoms is not reflected in this systematic review because of the search criteria focusing on MR-only clinical implementation pathways, rather than CT-MRI pathways. As such, further comment on their potential benefit in this context is not possible.

The anthropomorphic phantoms from Sun et al[25] and Cunningham et al[26] are a beneficial development in phantom design because they allow quantitative end-to-end testing, including dose measurements within the phantoms. Of interest is the ability of one set of phantoms[26] to model physiologic changes in the bladder and rectum, which will improve the commissioning process by allowing the effects of patient anatomic changes to be assessed in a quantitative and reproducible manner. Further development of anthropomorphic phantoms could increase their use in the quality assurance of MR-only planning as it develops as a clinical technique.

For MR-only planning to be implemented, MR data are required to be used for patient position verification purposes before treatment. Publications relating to data assessment as part of clinical implementation were limited, however, with only Tyagi et al,[18] Kemppainen et al,[33] and Korhonen et al[34] assessing prostate patient treatment positioning accuracy. All 3 studies assessed DRR positioning, with results showing broad agreement between the use of planning CT and synthetic CT−generated DRRs, providing confidence that synthetic CT−generated DRRs can produce clinically acceptable results. Korhonen et al[34] and Kemppainen et al[33] also investigated interobserver and intraobserver variability for DRR registrations and found it to be clinically insignificant.

Three-dimensional CBCT−based patient treatment positioning was investigated using manual fiducial marker registration[18] and the automatic bone and gray-value registration methods of Elekta's XVI system.[34] Synthetic CT−to-CBCT registrations were comparable to planning CT-to-CBCT registrations (mean differences <1 mm), indicating that synthetic CT data sets can replace CT data sets for manual or automatic registrations and for patient treatment positioning. Tyagi et al[18] also noted that, anecdotally, clinicians were happy with the delineations of bladder and rectum on the synthetic CT. It is interesting that the synthetic CT−to-CT results varied between Korhonen et al[34] and Tyagi et al.[18] This could be influenced by a number of factors, including the difference in matching technique (automatic vs manual) or the inclusion of a patient mold within Tyagi's study to improve setup reproducibility.

MR-to-CBCT registrations were also assessed by Korhonen et al,[34] but they did not replicate the same level of similarity to planning CT-to-CBCT registrations as synthetic CT−to-CBCT registrations. This result is understandable because XVI uses a chamfer matching algorithm and is optimized for registering data sets of the same modality (ie, CT to CBCT), and registrations could improve if more suitable mutual information algorithms were used. It

Table 5 Summary of the key results from the MR-only commissioning processes studies

Author	Year	No. of patients in study*	Other key information
Kapanen et al[35]	2013	X	Proposed calibration and testing procedures for verification of the treatment isocenter position, geometric accuracy, and other basic QA with an ACR phantom
Korsholm et al[37]	2014	21	A statistical model approach to assessing the accuracy of sCT calculation was used where the criteria of accuracy was considered to be 95% of patients having an uncertainty in dose calculation within the PTV within 2% of the CT dose
Maspero et al[21]	2017	14	For electron density conversion, sCT generation, and interscan difference, average dose difference in the CTV of 0.7% ± 0.2%, 0.16% ± 0.13%, and 0.01% ± 0.35% and in the whole body of 0.1% ± 0.03%, −0.03% ± 0.02%, and 0 ± 0.06% were found, respectively
Kerkmeijer et al[15]	2018	X	Recommended requirements for MR-only radiation therapy clinical implementation including geometric accuracy, treatment position MR acquisition, sCT generation, MRI-based OAR delineation, and protocol optimization and MRI-based treatment position verification
Kim et al[36]	2018	X	1. Many processes and therefore failure modes are shared between CT-MR and MR-only workflows with the highest failure modes related to changes in target location due to internal anatomy changes, in these cases current mitigation processes were still valid 2. The highest risk failure modes for the MR-only workflow alone related to the sCT generation process, including: inaccuracies in target delineation on MR images, insufficient management of patient- and system-level distortions and inaccurate bone volumes 3. Mitigation strategies for failures include sufficient staff training and a robust quality-control and quality-assurance program
Palmer et al[38]	2018	10	1. The CBCT system was stable over time in HU (standard deviation <40 HU) and the variation in HU between CT and CBCT was found to be minimal (<60 HU) 2. A comparison of the dose distributions between sCT and CT compared with sCT and CBCT found mean dose differences for all metrics of ≤1% 3. The CBCT system can be considered to be similar to a CT system and can be used as a clinically feasible QA procedure

Abbreviations: ACR = American College of Radiology; CBCT = cone beam CT; CT = computed tomography; HU = Hounsfield units; MR = magnetic resonance; MRI = magnetic resonance imaging; PTV = planning target volume; QA = quality assurance; sCT = synthetic CT.
* "No. patients in Study" refers to the total number of patients recruited for the MR-only investigations.

is an indicator that CT cannot be simply replaced with MRI and that further commercial support and investment in this field is required. In addition, it is important to note that although differences are seen between CT and MR registrations, from the data shown, neither CT nor MR can be determined as more accurate because there is no ground truth for comparison. It can only be determined that the registrations produce different results. To resolve which modality is more accurate, manual landmark evaluation can be used for an initial comparison, and a future potential solution would be to use an anthropomorphic phantom that could provide the required ground truth information. It can be hypothesized that it would be best to register MR to CBCT, rather than synthetic CT to CBCT, because this would mean that real rather than synthesized data were being used, thus theoretically improving the registration accuracy. These findings suggest that MR-only pathways exist that allow reproducible patient positioning verification to be completed. In addition, these studies provide a suitable methodology for a center looking to implement MR-only planning with respect to the assessment of patient treatment positioning accuracy and reproducibility.

A wide variety of processes and experiences relating to commissioning an MR-only pathway were reported for prostate treatments. The experiences of Kerkmeijer et al[15] and Kapanen et al[35] in terms of workflow, equipment, and commissioning requirements provide substantial amounts of information, which is particularly beneficial because these processes, within this early phase of clinical implementation, are not well established. It is a challenging process to determine the commissioning and routine workflow to ensure optimal performance and the highest quality of patient care; therefore, more publications detailing individual centers' experiences, such as these, would be welcome until this technique is more firmly embedded in routine clinical practice or guidance documents are published. The information provided within the work of Kim et al[36] provides useful tools for identifying risks and highlights many risks that will be shared among all MR-only pathways; it also suggests mitigation strategies to lessen their influence. That the greatest source of risk is the synthetic CT generation process is not a surprising result; however, the strength of this methodology is that it provides an overall framework for assessing, comparing, and minimizing risks. This methodology also allows the user to have confidence that their MR-only pathway is optimized to protect from errors as much as is reasonably practicable. In addition, as a process, it can be repeated over time to

reassess and fine-tune the pathway continually; it is also applicable to any future MR-only treatment sites in addition to prostate cancer, which was presented here.

The quantification of the accuracy of a center's local synthetic CT generation technique is a key stage of commissioning an MR-only pathway, and the studies by Maspero et al[21] and Korsholm et al[37] provide differing methods of undertaking this. It is interesting to see that Maspero et al[21] found that electron density conversion (from CT scan−generated Hounsfield units to electron density via an election density plotted curve) was the greatest confounding factor, followed by synthetic CT generation (the assigning of Hounsfield units to MR scan voxels to produce the synthetic CT). Interscan differences (setup and positioning differences between CT and MRI scans, MR geometric inaccuracy, and CT-MRI registration errors as required for comparison) produced almost a negligible difference in result. This finding suggests that the commissioning process should also focus on appropriate electron density curve calibration as a key part of the commissioning process.

Palmer et al[38] present a method of validating synthetic CT generations using collected patient CBCT data and provide a tool by which commissioning centers can ensure further confidence over the accuracy of their treatment planning pathway. As previously discussed by Kim et al,[36] the generation of synthetic CTs is a major risk in the MR-only pathway, and the challenge of ensuring robust patient treatment on an individual basis is nontrivial. The method of Palmer et al would directly allow a gross error check on the treatment plan, which could highlight potential issues at the beginning of a patient's treatment. Palmer et al[38] noted, however, that further analysis of this technique would be beneficial, as only simple errors were assessed within the validation presented; as a consequence, its sensitivity to less-noticeable errors is uncertain.

The studies identified here are a significant step toward widespread pelvic MR-only clinical implementation; however, further attention is required in several areas. As discussed previously, the translation of this technique to other clinical pelvic sites is a significant challenge that should not be underestimated. Several studies reported issues associated with processing data within MR-only pathways. These issues may be due to a lack of support for the MR-only workflow by radiation therapy vendors. Clinical treatment planning software and further collaboration, investment, or support from commercial companies would be beneficial. Unseen in the literature was a long timescale ($\geq$1 year) Bo distortion evaluation study focusing on its effects on an MR-only pathway. Long timescale changes in Bo distortion could have a significant influence on resultant MRI geometric accuracy, which would require correction to prevent the translation of errors into the planning process. Such a study would demonstrate the reliability and reproducibility of scanner Bo distortion over time and therefore provide evidence for distortion quality assurance frequency recommendations. Patient treatment positioning verification within clinical implementation was addressed in only 3 studies. Although the results presented were encouraging and suggested that MR-only techniques can accurately be used for patient treatment positioning, there is plenty of evidence yet to be gathered. All these studies involved small cohorts of patients and used the synthetic CT generation method of the Philips prostate MRCAT[11] or similar,[39] which are not comparable to all methods of synthetic CT generation. In addition, the majority of results were collected with manual registration techniques, whereas it is common in the clinic to use a manufacturer's automated or semiautomated technique. The effects different clinical equipment and techniques, in larger patient cohorts, need to be investigated fully.

Despite the variety of publications related to MR-only commissioning and individual centers experiences, the radiation therapy community is so varied in term of equipment, resources, and technique that there is significant scope for further experience to be reported in the literature and consensus guidelines to be produced by early adopters and national bodies. There is also a substantial need for more studies to begin providing evidence of the benefit of using MR-only planning, such as improved patient outcomes or treatment pathway improvements.

Conclusion

MR-only planning has been clinically implemented for the treatment of prostate cancer; however, further research is required to develop MR-only planning for other pelvic sites. In particular, the accuracy of synthetic CT generation models for female anatomy requires further reporting within the literature. MR scanner distortions are no longer a barrier to MR-only planning, although they must be managed appropriately, whereas MR data acquisition and synthetic CT generation for prostate treatments have been shown to be sufficiently accurate for clinical use. The clinical implementation of MR-only patient treatment positioning verification remains underreported in the literature and requires substantial investigation to allow its widespread use. The range of investigations reported here are a suitable starting point for radiation therapy centers aiming to clinically implement MR-only planning; however, further evidence and regulation is required, including the publication of consensus guidelines from early adopters and governing bodies.

References

1. Njeh CF. Tumor delineation: The weakest link in the search for accuracy in radiotherapy. *J Med Phys* 2008;33:136-140.
2. Sander L, Langkilde NC, Holmberg M, Carl J. MRI target delineation may reduce long-term toxicity after prostate radiotherapy. *Acta Oncol* 2014;53:809-814.
3. Wortel RC, Heemsbergen WD, Smeenk RJ, et al. Local Protocol Variations for Image Guided Radiation Therapy in the Multicenter Dutch Hypofractionation (HYPRO) Trial: Impact of rectal balloon and

MRI delineation on anorectal dose and gastrointestinal toxicity levels. *Int J Radiat Oncol Biol Phys* 2017;99:1243-1252.
4. Johnstone E, Wyatt JJ, Henry AM, et al. Systematic review of synthetic computed tomography generation methodologies for use in magnetic resonance imaging-only radiation therapy. *Int J Radiat Oncol Biol Phys* 2018;100:199-217.
5. Owrangi AM, Greer PB, Glide-Hurst CK. MRI-only treatment planning: Benefits and challenges. *Phys Med Biol* 2018;63:05TR01.
6. Persson E, Gustafsson C, Nordström F, et al. MR-OPERA: A multicenter/multivendor validation of magnetic resonance imaging-only prostate treatment planning using synthetic computed tomography images. *Int J Radiat Oncol Biol Phys* 2017;99:692-700.
7. Karlsson M, Karlsson MG, Nyholm T, Amies C, Zackrisson B. Dedicated magnetic resonance imaging in the radiotherapy clinic. *Int J Radiat Oncol Biol Phys* 2009;74:644-651.
8. Weygand J, Fuller CD, Ibbott GS, et al. Spatial precision in magnetic resonance imaging-guided radiation therapy: The role of geometric distortion. *Int J Radiat Oncol Biol Phys* 2016;95:1304-1316.
9. Edmund JM, Nyholm T. A review of substitute CT generation for MRI-only radiation therapy. *Radiat Oncol* 2017;12:26; [Review].
10. Siversson C, Nordström F, Nilsson T, et al. Technical note: MRI only prostate radiotherapy planning using the statistical decomposition algorithm. *Med Phys* 2015;42:6090-6097.
11. Köhler M, Vaara T, Van Grootel M, et al. MR-Only Simulation for Radiotherapy Planning [white paper]. Amsterdam: Koninklijke Philips; 2015.
12. Philips Healthcare. MR-only simulation pelvis 2018. Available at: https://www.philips.co.uk/healthcare/product/HCNMRF171/mr-only-simulation-unleash-the-real-power-of-mr-simulation. Accessed June 2, 2019.
13. Moher D, Liberati A, Tetzlaff J, Altman DG, and PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 2009;6:e1000097.
14. Tenhunen M, Korhonen J, Kapanen M, et al. MRI-only based radiation therapy of prostate cancer: Workflow and early clinical experience. *Acta Oncol* 2018;57:902-907.
15. Kerkmeijer LGW, Maspero M, Meijer GJ, et al. Magnetic resonance imaging only workflow for radiotherapy simulation and planning in prostate cancer. *Clin Oncol* 2018;30:692-701.
16. Christiansen RL, Jensen HR, Brink C. Magnetic resonance only workflow and validation of dose calculations for radiotherapy of prostate cancer. *Acta Oncol* 2017;56:787-791.
17. Tyagi N, Fontenla S, Zelefsky MM, et al. Clinical workflow for MR-only simulation and planning in prostate. *Radiat Oncol* 2017;12:119.
18. Tyagi N, Fontenla S, Zhang J, et al. Dosimetric and workflow evaluation of first commercial synthetic CT software for clinical use in pelvis. *Phys Med Biol* 2017;62:2961-2975.
19. Kemppainen R, Suilamo S, Tuokkola T, et al. Magnetic resonance-only simulation and dose calculation in external beam radiation therapy: A feasibility study for pelvic cancers. *Acta Oncol* 2017;56:792-798.
20. Wyatt JJ, Dowling JA, Kelly CG, et al. Investigating the generalisation of an atlas-based synthetic-CT algorithm to another centre and MR scanner for prostate MR-only radiotherapy. *Phys Med Biol* 2017;62:N548-N560.
21. Maspero M, Seevinck PR, Schubert G, et al. Quantification of confounding factors in MRI-based dose calculations as applied to prostate IMRT. *Phys Med Biol* 2017;62:948-965.
22. Price RG, Knight RA, Hwang KP, et al. Optimization of a novel large field of view distortion phantom for MR-only treatment planning. *J Appl Clin Med Phys* 2017;18:51-61.
23. Huang KC, et al. Phantom-based characterization of distortion on a magnetic resonance imaging simulator for radiation oncology. *Phys Med Biol* 2016;61:774-790.
24. Walker A, Cao Y, Baharom U, Balter JM. Continuous table acquisition MRI for radiotherapy treatment planning: Distortion assessment with a new extended 3D volumetric phantom. *Med Phys* 2015;42:1982-1991.
25. Sun J, Dowling J, Pichler P, et al. MRI simulation: End-to-end testing for prostate radiation therapy using geometric pelvic MRI phantoms. *Phys Med Biol* 2015;60:3097-3109.
26. Cunningham JM, Barberi EA, Miller J, et al. Development and evaluation of a novel MR-compatible pelvic end-to-end phantom. *J Appl Clin Med Phys* 2019;20:265-275.
27. Gustafsson C, Nordström F, Persson E, Brynolfsson J, Olsson LE. Assessment of dosimetric impact of system specific geometric distortion in an MRI only based radiotherapy workflow for prostate. *Phys Med Biol* 2017;62:2976-2989.
28. Glide-Hurst C, Nejad-Davarani S, Weiss S, et al. Per-organ assessment of subject-induced susceptibility distortion for MR-only male pelvis treatment planning. *Radiat Oncol* 2018;13:149.
29. Adjeiwaah M, Bylund M, Lundman JA, et al. Quantifying the effect of 3T magnetic resonance imaging residual system distortions and patient-induced susceptibility distortions on radiation therapy treatment planning for prostate cancer. *Int J Radiat Oncol Biol Phys* 2018;100:317-324.
30. Wyatt J, Hedley S, Johnstone E, et al. Evaluating the repeatability and set-up sensitivity of a large field of view distortion phantom and software for magnetic resonance-only radiotherapy. *Phys Imaging Radiat Oncol* 2018;6:31-38.
31. Torfeh T, Hammoud R, Perkins G, et al. Characterization of 3D geometric distortion of magnetic resonance imaging scanners commissioned for radiation therapy planning. *Magn Reson Imaging* 2016;34:645-653.
32. Price RG, Kadbi M, Kim J, et al. Technical note: Characterization and correction of gradient nonlinearity induced distortion on a 1.0 T open bore MR-SIM. *Med Phys* 2015;42:5955-5960.
33. Kemppainen R, Vaara T, Joensuu T, Kiljunen T. Accuracy and precision of patient positioning for pelvic MR-only radiation therapy using digitally reconstructed radiographs. *Phys Med Biol* 2018;63:055009.
34. Korhonen J, Kapanen M, Sonke JJ, et al. Feasibility of MRI-based reference images for image-guided radiotherapy of the pelvis with either cone-beam computed tomography or planar localization images. *Acta Oncol* 2015;54:889-895.
35. Kapanen M, Collan J, Beule A, et al. Commissioning of MRI-only based treatment planning procedure for external beam radiotherapy of prostate. *Magn Reson Med* 2013;70:127-135.
36. Kim J, Miller B, Siddiqui MS, et al. FMEA of MR-only treatment planning in the pelvis. *Adv Radiat Oncol* 2018;4:168-176.
37. Korsholm ME, Waring LW, Edmund JM. A criterion for the reliable use of MRI-only radiotherapy. *Radiat Oncol* 2014;9:16.
38. Palmer E, Persson E, Ambolt P, et al. Cone beam CT for QA of synthetic CT in MRI only for prostate patients. *J Appl Clin Med Phys* 2018;19:44-52.
39. Korhonen J, Kapanen M, Keyriläinen J, Seppälä T, Tenhunen M. A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer. *Med Phys* 2014;41:011704.
40. Dowling JA, Sun J, Pichler P, et al. Automatic substitute computed tomography generation and contouring for magnetic resonance imaging (MRI)-alone external beam radiation therapy from standard MRI sequences. *Int J Radiat Oncol Biol Phys* 2015;93:1144-1153.