Prostate Imaging Quality (PI-QUAL): a new quality control scoring system for multiparametric MRI of the prostate from the PRECISION trial

Francesco Giganti [1,2], Clare Allen [1], Mark Emberton [2,3],

*Caroline M Moore [2,3] *, Veeru Kasivisvanathan [2,3] * for the PRECISION study group*

* These authors share joint senior authorship

1. Department of Radiology, University College London Hospital NHS Foundation Trust, London, UK
2. Division of Surgery & Interventional Science, University College London, London, UK
3. Department of Urology, University College London Hospital NHS Foundation Trust, London, UK

Corresponding author:

Francesco Giganti, MD
Division of Surgery and Interventional Science,
University College London, 3rd Floor,
Charles Bell House, 43-45 Foley St.,
London, United Kingdom
W1W 7TS
email: f.giganti@ucl.ac.uk

Abstract

The PRECISION trial was a multi-centre randomised study which demonstrated that mpMRI-targeted biopsy was superior to standard transrectal ultrasound-guided biopsy for the detection of prostate cancer. The outcomes of studies reporting mpMRI-targeted biopsies are dependent on the quality of the mpMRI but there are currently no scoring systems available for evaluating this. We introduced a novel scoring system, the Prostate Imaging Quality (PI-QUAL) score, to assess the quality of scans in the PRECISION trial. PI-QUAL is a 1-to-5 Likert score, where 1 means that no mpMRI sequences are of diagnostic quality and 5 implies that each sequence is independently of optimal diagnostic quality. Fifty-eight out of 252 (23%) of the mpMRI scans, chosen at random from each of the 22 centres in this trial were evaluated by two experienced radiologists from the coordinating trial centre, in consensus, blinded to pathology results. Overall, the mpMRI quality in the centres participating in PRECISION was appropriate. MpMRI quality was of sufficient diagnostic quality (PI-QUAL ≥ 3) for 55 (95%) scans, with 35 (60%) being of adequate or optimal diagnostic quality (PI-QUAL ≥ 4). The mpMRI scan quality was highest for T2-weighted imaging and lowest for the dynamic contrast-enhanced acquisitions. Further validation of this scoring system is warranted.

Patient summary:

In this study we developed a scoring system (PI-QUAL) to assess the quality of multiparametric magnetic resonance imaging (mpMRI) in prostate cancer detection. We used the scans from 22 centres that participated in the PRECISION trial. Though there was room for improvement on the images which used intravenous contrast, we found that mpMRI in the PRECISION trial was of sufficient diagnostic quality (i.e. (PI-QUAL ≥ 3) for 95% of the scans.

Text

A key determinant of the likelihood of being able to detect and rule out clinically significant prostate cancer (PCa) on multiparametric magnetic resonance imaging (mpMRI) is the quality of the scan. However, there are currently no scoring systems available for evaluating this. This has implications for being able to interpret any study of mpMRI and mpMRI-targeted biopsy, as well as for enabling men to safely avoid unnecessary biopsies. A quality assessment is important at both centre and scanner level, where scan optimisation is required, and at patient level where both transient and modifiable factors (e.g. movement, rectal gas) and permanent factors (e.g. hip replacement) may reduce scan quality. The importance of prostate mpMRI quality was recently advocated by a panel of international experts following a Delphi consensus process. [1]

The diagnostic quality of the scan is determined by how adequate each dominant sequence [i.e. T2-weighted imaging (T2-WI) for the transition zone and diffusion weighted imaging (DWI) for the peripheral zone, in addition to dynamic contrast enhancement (DCE) when it comes to discriminate equivocal lesions in the peripheral zone] is for being able to rule in or rule out clinically significant PCa.

The PRECISION trial [2] was a multi-centre randomised study that showed the superiority of mpMRI-targeted to standard transrectal ultrasound-guided biopsy in 500 biopsy-naïve men. Men were scanned using either a 3T or 1.5T scanner, and three centres used an endorectal coil [2]. Full details of the mpMRI scanners used were reported in the Supplementary Appendix of the PRECISION study [2].

During the quality assurance work in PRECISION, a randomisation process was performed by computer-generated random sequences, leading to selection of 25% of each centre's scans. The scans were reviewed in consensus at the coordinating centre by two expert radiologists who actively participate at our weekly prostate multidisciplinary meetings (FG and CA, with 7 and 20 years of experience in prostate mpMRI respectively, and reporting ≥ 1,800 and ≥ 3,000 prostate mpMRI scans per year, repsectively) and who used objective criteria within the scan that could be used to assess quality [3].
The radiologists were blinded to the local radiologist's reports and to biopsy results. Six scans from one centre were excluded as they could not provide the full contrast sequence for central review.

We developed a 1-to-5 scoring system, called the Prostate Imaging Quality (PI-QUAL) score (Table 1) and evaluated all scans according to this. The PI-QUAL score indicates the adequacy of the diagnostic quality of an mpMRI scan, where 1 means all sequences are below the minimum standard of diagnostic quality, 3 means that the scan is of sufficient diagnostic quality and 5 means that all three sequences are of optimal diagnostic quality.

The PI-QUAL score is derived by evaluating the mpMRI against a defined set of objective quality criteria in line with PI-RADS v.2 guidelines [3] that covered adequacy of the conduct of all three of the mpMRI sequences (Fig. 1), as PI-RADS v. 2.1 guidelines [4] had not been published at the time of PRECISION. Additional objective criteria were introduced (e.g. the presence of artefacts, and ability to delineate the capsular vessels) all of which were considered to be consistent with a high-quality scan. See Figure 2 for examples of different PI-QUAL scores. Each sequence was given a summary binary outcome for whether or not the sequence was of diagnostic quality and an overall final PI-QUAL score was given. There is a subtle but important difference between each PI-QUAL score. In particular, a PI-QUAL score of 4 is different from a PI-QUAL score of 5 in that for this latter each sequence is of optimal diagnostic quality (i.e. only for PI-QUAL 5 all acquisition parameters are fully compliant with PI-RADS v. 2 guidelines from a technical point of view) [3].

Overall, the mpMRI quality in the centres participating in PRECISION was appropriate. MpMRI quality was of at least sufficient diagnostic quality (PI-QUAL ≥ 3) for 55 (95%) scans, with 35 (60%) being of adequate or excellent quality (PI-QUAL ≥ 4). Twelve scans (21%) had a score of 5, 23 (40%) had a score of 4, 20 scans (34%) had a score of 3 and three scans had a score < 3. The diagnostic quality for each sequence was highest for T2-WI (55/58; 95%), then for DWI (46/58; 79%) and lowest for DCE (38/58; 66%).

There was very good adherence to PI-QUAL quality criteria for T2-WI. The position of the axial plane can vary between Institutions (e.g. perpendicular to the MR table, orthogonal to the rectum [5] or in an oblique axial plane matching the long axis of the prostate, as shown in Supplementary Fig. 1) but what is crucial is that the other sequences (i.e. DWI and DCE) are acquired in the same plane. We noted that only three scans had a PI-QUAL < 3, and the reason was that these scans had inadequate in-plane and spatial resolution (n=1) or severe motion artefacts (n=2).

For DWI, we noted that 34/58 (59%) scans were lacking a dedicated single high b-value acquisition, even though the PI-RADS v.2 guidelines recommend at least a minimum b value of 1,400 s/mm^2 at 1.5T (but up to 2,000 s/mm^2, if adequate signal-to-noise ratio permits). We also observed that 7/58 (12%) scans had a synthesized (i.e. computer-generated b values) apparent diffusion coefficient map rather than an acquired one, and overall DWI was of not of diagnostic quality for these scans.

The highest variability in quality of mpMRI sequences was for DCE. PI-RADS v.2 guidelines recommended a temporal resolution < 10 seconds (preferably < 7 seconds) and we noted that only 31/58 (53%) scans matched the 10 seconds cut off. However, we know from PI-RADS v. 2.1 guidelines [4], which were produced after the PRECISION study, that temporal resolution can be increased to 15 seconds in order to achieve higher spatial resolution, and 21/23 (91%) centres participating in PRECISION had a temporal resolution ≤ 15 seconds. We

also know that fat suppression is very important to optimise contrast to noise ratio but 26/58 (45%) scans in this study did not have subtracted or fat suppressed DCE images. The role of DCE has been recently downplayed in the PI-RADS v.2.1 guidelines but we also know from a consensus meeting on prostate mpMRI [5] that DCE should be still used as a "safety net" when the quality of T2-WI or DWI is inadequate.

Supplementary Table 1 shows the detection rate for clinically significant cancer (defined as Gleason ≥ 3+4) in 38/58 (66%) men who received a biopsy after mpMRI as a function of the PI-QUAL score.

We know that high-quality prostate mpMRI is critical, especially before biopsy and in order to defer/avoid one. Despite some variation from the technical standards and image quality across site, our results suggest that the mpMRI quality in the centres participating in PRECISION was appropriate. The PRECISION data allow us to conclude that an mpMRI-led prostate cancer diagnostic pathway including targeted biopsies can outperform standard biopsy when the quality of mpMRI acquisition is adequate. There is room for improvement, particularly in the assessment of T2-WI (i.e. should each T2-WI sequence be scored separately or should all T2-WI sequences be treated as a whole, as we did?) and in the quality of the DCE sequences, which may influence the value of PI-QUAL in the detection of clinically significant PCa. Future discussions of a detailed flowchart of how the PI-QUAL score is attributed, taking into account all possible combinations of different sequences, are warranted.

Because there is no mpMRI prostate phantom, we chose structures within the prostate that we thought should be visible, and these were used as an objective marker of the quality of that particular mpMRI sequence (especially for T2-WI and DCE sequences). This does however give room for refinements in subsequent studies.

A possible limitation of this preliminary report is that PRECISION was an international multi-centre study that involved academic and non-academic centres, with different levels of experience in prostate biopsies. However, the level of experience of the radiologists and the quality of mpMRI involved in the study might have been higher than the average standard. Nevertheless, PI-QUAL is derived from an international set of images, in order to have an initial general idea on how different centres with different levels of expertise perform prostate mpMRI across the world (UK, Europe, North and South America). Our report combines robust data from an international multi-centre randomised trial (PRECISION) and the trial's extensive multicentre dataset on prostate mpMRI, which is a unique and valuable resource.

However, it should be kept in mind that PI-QUAL represents the basis for future works, and though the PI-QUAL scoring system has been evaluated in this large international multi-centre study, it requires validation and optimisation in future studies. Broad consensus among experts should be reached before using this scoring system in clinical settings. Some

image features will need to be refined; for example, which axial plane acquisition should be defined as 'optimal' or on how the quality of the dominant sequence should be weighted in the final assessment of the score. In addition to this, it is necessary to clarify whether synthesised and acquired high b values acquisitions should be weighted equally, and if the quality of the ADC map is linked to this.

We believe that PI-QUAL is the first attempt at producing an mpMRI quality scoring system with its strengths that it arises from the experience of an international multicentre trial in prostate mpMRI. For the time being, PI-QUAL offers clinicians the only available scoring system of evaluating and reporting quality of their prostate mpMRI so that the generalisability of results from a study can be evaluated. We plan to undertake the external validation of PI-QUAL, in the hope of corroborating these initial results. We anticipate that the composition of the scoring system will undergo further refinements based on this.

Table 1: Assessment of the diagnostic quality of multiparametric MRI scans (PI-QUAL score)
Legend: PI-QUAL: Prostate Imaging QUALity; mpMRI: multiparametric magnetic resonance imaging; PI-RADS: Prostate Imaging–Reporting and Data System

PI-QUAL score	Criteria	Clinical Implication
1	All mpMRI sequences are below the minimum standard of diagnostic quality	It is NOT possible to rule in all significant lesions *
2	Only one mpMRI sequence is of acceptable diagnostic quality	It is NOT possible to rule out all significant lesions *
3	At least two mpMRI sequences taken together are of diagnostic quality	It is possible to rule in all significant lesions It is NOT possible to rule out all significant lesions
4	Two or more mpMRI sequences are independently of diagnostic quality	It is possible to rule in all significant lesions
5	All mpMRI sequences are of optimal diagnostic quality	It is possible to rule out all significant lesions elsewhere

* Reports should not include PI-RADS or Likert scores.

Legend: PI-QUAL: Prostate Imaging QUALity; mpMRI: multiparametric magnetic resonance imaging; PI-RADS: Prostate Imaging Reporting and Data System

Fig. 1: Scoring sheet to assess the quality of multiparametric MRI using the PI-QUAL scoring system.

Scan & site number:

⊙ PRECISION

Prostate Imaging QUALity control (PI-QUAL) scoring sheet

PI-QUAL score	Criteria	Clinical Implication
1	All MR sequences are below the minimum standard of diagnostic quality	It is NOT possible to rule in all significant lesions [§]
2	Only one MR sequence is of acceptable diagnostic quality	It is NOT possible to rule out all significant lesions [§]
3	At least two sequences taken together are of diagnostic quality	It is possible to rule in all significant lesions It is NOT possible to rule out all significant lesions
4	Two or more sequences are independently of diagnostic quality	It is possible to rule in all significant lesions
5	All sequences are of optimal diagnostic quality	It is possible to rule out all significant lesions elsewhere

[§] Reports should not include PI-RADS or Likert scores

Please (✓) if present *:

T2-WI	DWI	DCE
Axial plane	Adequate field of view	Adequate field of view
Sagittal plane	Adequate in-plane resolution	Adequate in-plane resolution
Coronal plane	Adequate slice thickness	Adequate slice thickness
Adequate field of view	Multiple *b* values acquired	Pre-contrast T1-WI available
Adequate in-plane resolution	Dedicated high *b* value sequence	Fat suppression/saturation
Adequate slice thickness	ADC map of good quality	Post-contrast T1-WI available
Z-axis correctly positioned	Derived ADC map	Adequate temporal resolution
Anatomical resolution	Rectal air artefact	Adequate acquisition duration
Capsule delineation		Delineation of the capsular vessels
Inclusion of seminal vesicles		Delineation of the vessels in the adductor canal
Inclusion of ejaculatory ducts		
Neurovascular bundles clearly delineated		
Sphincter muscle clearly delineated		
Motion artefacts		

Is T2-WI of diagnostic quality?
☐ Yes
☐ No

Is DWI of diagnostic quality?
☐ Yes
☐ No

Is DCE of diagnostic quality?
☐ Yes
☐ No

PI-QUAL score:
1 ☐
2 ☐
3 ☐
4 ☐
5 ☐
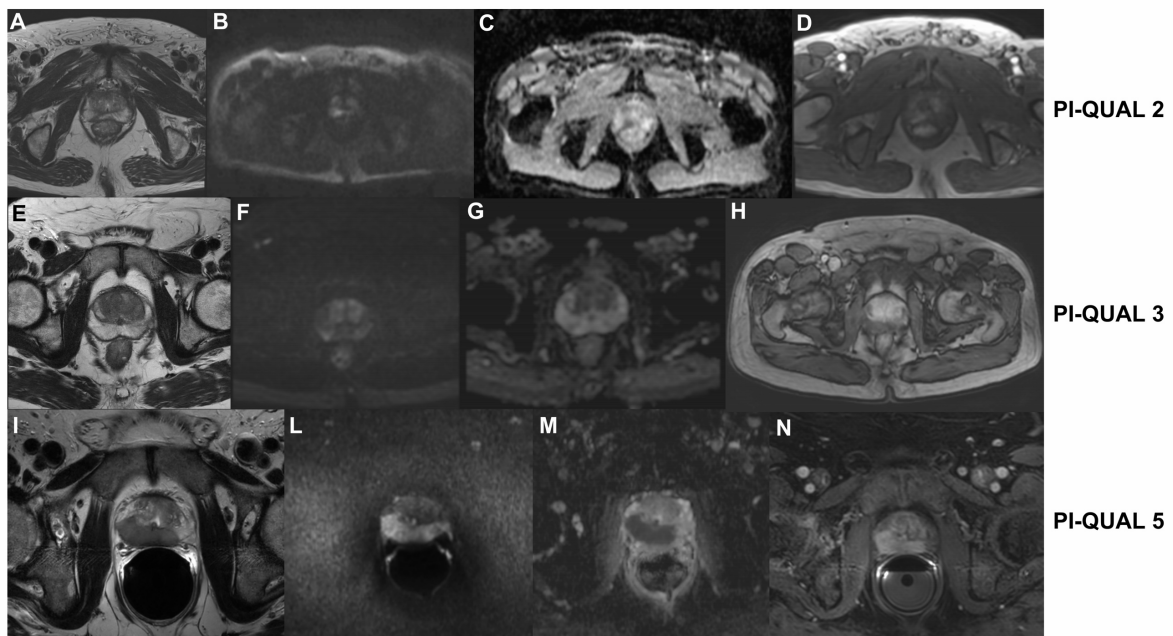
Comments:

Date:

Reporting Radiologist:

Signed:

* According to PI-RADS v. 2 guidelines (2015)

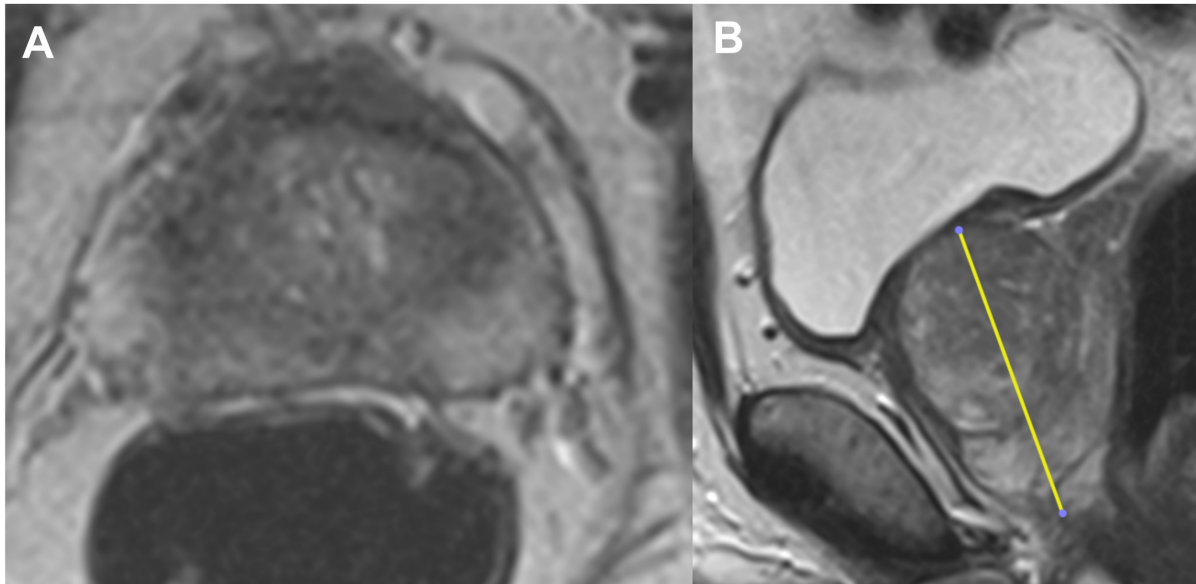**Fig. 2**: Three examples of different PI-QUAL scores.

(A-D): The image acquisition on axial T2-weighted imaging (A) is adequate (field of view: 17x20cm; PI-RADS v.2.0 recommendation: 12-20 cm) with a sagittal acquisition straight axial to the patient (not shown). The field of view of DWI is too large (field of view: 34x26 cm; PI-RADS v.2.0 recommendation: 16-22 cm) (B-C) and there is no dedicated high b sequence. There is no fat suppression on dynamic contrast enhanced sequences (C), with no clear definition of the capsular vessels and the temporal resolution is 12 seconds (PI-RADS v.2.0 recommendation: ≤ 10 seconds). As only T2-weighted imaging is independently of acceptable diagnostic quality, the final PI-QUAL score is 2.

(E-H): The image acquisition on axial T2-weighted imaging (E) is adequate (field of view: 18x18cm; PI-RADS v.2.0 recommendation: 12-20 cm) with a sagittal acquisition straight axial to the patient (not shown). DWI is adequate (field of view: 18x20 cm; PI-RADS v.2.0 recommendation: 16-22 cm) (F-G), with multiple b values (0, 800, 1600 s/mm^2 but there is no dedicated high b sequence. There is no fat suppression on dynamic contrast enhanced sequences (H), with no clear definition of the capsular vessels, despite an adequate temporal resolution (3 seconds; PI-RADS v.2.0 recommendation: ≤ 10 seconds). As two sequences taken together (T2-WI and DWI) are of diagnostic quality, the final PI-QUAL score is 3.

(I-N): The image acquisition on axial T2-weighted (I) is adequate (field of view: 18x18cm; PI-RADS v.2.0 recommendation: 12-20 cm) with additional sagittal and coronal acquisitions (not shown). Please note the correct positioning of the endorectal coil and the presence of some artefacts does not prevent from assessing the whole prostate. DWI is adequate (field of view: 18x18 cm; PI-RADS v.2.0 recommendation: 16-22 cm) (L-M), with a dedicated high b value sequence (1,500 s/mm^2; PI-RADS v.2.0 recommendation: 1,400-2,000 s/mm^2) (L). There is fat suppression on dynamic contrast enhanced sequences (N), with an adequate temporal resolution of 8 seconds (PI-RADS v.2.0 recommendation: ≤ 10 seconds). As all sequences are of optimal diagnostic quality, the final PI-QUAL score is 5.

Suppl. Fig. 1: Axial (A) and sagittal (B) planes of prostate mpMRI, showing the long axis of the prostate from base to apex (yellow vertical line in B).

Supplementary Table 1: detection rate for clinically significant cancer (defined as Gleason ≥ 3+4) according to PI-QUAL score

PI-QUAL score	Significant cancer identified (n)	No significant cancer identified (n)
1 (n=1)	0	1
2 (n=2)	1	1
3 (n=13)	10	3
4 (n=14)	6	8
5 (n=8)	4	4

Legend – PI-QUAL: Prostate Imaging Quality

References

[1]   De Rooij M, Israël B, Tummers M, et al. ESUR/ESUI consensus statements on multi-
        parametric MRI for the detection of clinically significant prostate cancer: quality
        requirements for image acquisition, interpretation and radiologists ' training EEur
        Radiol 2020 (in press); doi: 10.1007/s00330-020-06929-z

[2]   Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-Targeted or Standard Biopsy for
        Prostate-Cancer Diagnosis. N Engl J Med 2018;378(19):1767–1777.

[3]   Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging - Reporting and
        Data System: 2015, Version 2. Eur Urol 2016;69(1):16–40.

[4]   Turkbey B, Rosenkrantz AB, Haider MA, et al. Prostate Imaging Reporting and Data
        System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System
        Version 2. Eur Urol 2019;76(3):340-351.

[5]   Engels RRM, Israel B, Padhani AR, Barentsz JO. Multiparametric Magnetic Resonance
        Imaging for the detection of clinically significant prostate cancer: what urologists
        need to know. Part 1: Acquisition. Eur Urol 2020;77:457–468.

[6]   Brizmohun Appayya M, Adshead J, Ahmed H, et al. National Implementation of multi-
        parametric MRI for prostate cancer detection – Recommendations from a UK
        consensus meeting. BJU Int 2018;122(1):13-25.