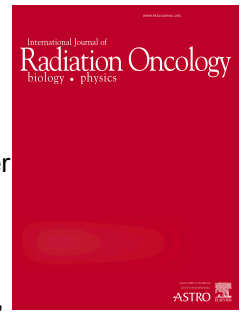
Accepted Manuscript



MRE11 as a predictive biomarker of outcome following radiotherapy in bladder cancer

Alexandra K. Walker, PhD, Katalin Karaszi, MSc, Helen Valentine, MRSC, Victoria Y. Strauss, PhD, Ananya Choudhury, MA PhD MRCP FRCR, Shaun McGill, BM BCh, Kaisheng Wen, MD, Michael D. Brown, PhD, Vijay Ramani, MB BS, MS, FRCS(Urol), Selina Bhattarai, MD, FRCPath, Mark TW. Teo, PhD, FRCR, Lingjian Yang, PhD, Kevin A. Myers, PhD, Nayneeta Deshmukh, MB ChB, FRCPath, Helen Denley, MB ChB, FRCPath, Lisa Browning, MB BS FRCPath, Sharon B. Love, BSc, Gopa Iyer, MD, Noel W. Clarke, MB BS, ChM, FRCS(Urol), Emma Hall, PhD, Robert Huddart, MA, MB BS, PhD, MRCP, FRCR, Nicholas D. James, MBBS, PhD, FRCP, FRCR, Peter J. Hoskin, MD, FRCP, FRCR, Catharine ML. West, PhD, Anne E. Kiltie, MA, DM, DSc, MRCP, FRCR

PII: S0360-3016(19)30371-2

DOI: https://doi.org/10.1016/j.ijrobp.2019.03.015

Reference: ROB 25597

To appear in: *International Journal of Radiation Oncology • Biology • Physics*

Received Date: 6 December 2018

Revised Date: 6 March 2019

Accepted Date: 9 March 2019

Please cite this article as: Walker AK, Karaszi K, Valentine H, Strauss VY, Choudhury A, McGill S, Wen K, Brown MD, Ramani V, Bhattarai S, Teo MT, Yang L, Myers KA, Deshmukh N, Denley H, Browning L, Love SB, Iyer G, Clarke NW, Hall E, Huddart R, James ND, Hoskin PJ, West CM, Kiltie AE, MRE11 as a predictive biomarker of outcome following radiotherapy in bladder cancer, *International Journal of Radiation Oncology • Biology • Physics* (2019), doi: https://doi.org/10.1016/j.ijrobp.2019.03.015.

MRE11 as a predictive biomarker of outcome following radiotherapy in bladder cancer

Alexandra K Walker, PhD[1]*, Katalin Karaszi, MSc[1]*, Helen Valentine, MRSC[2], Victoria Y Strauss, PhD[3†], Ananya Choudhury, MA PhD MRCP FRCR [2], Shaun McGill, BM BCh[1], Kaisheng Wen, MD[4], Michael D Brown, PhD[5], Vijay Ramani, MB BS, MS, FRCS(Urol)[6], Selina Bhattarai, MD, FRCPath[7], Mark TW Teo, PhD, FRCR[8], Lingjian Yang, PhD[2], Kevin A Myers, PhD[9], Nayneeta Deshmukh, MB ChB, FRCPath[4], Helen Denley, MB ChB, FRCPath[10], Lisa Browning, MB BS FRCPath[11,12], Sharon B Love, BSc[3], Gopa Iyer, MD[13], Noel W Clarke, MB BS, ChM, FRCS(Urol) [5,6], Emma Hall, PhD[14], Robert Huddart, MA, MB BS, PhD, MRCP, FRCR[15], Nicholas D James, MBBS, PhD, FRCP, FRCR[4], Peter J Hoskin, MD, FRCP, FRCR[16,17], Catharine ML West, PhD[2††], Anne E Kiltie, MA, DM, DSc, MRCP, FRCR[1††]

*These authors contributed equally to the work.

†This author was responsible for statistical analysis.

††These authors contributed equally to the work.

[1]CRUK/MRC Oxford Institute for Radiation Oncology, University of Oxford, Oxford OX3 7DQ, UK.

[2]Translational Radiobiology Group, Division of Cancer Sciences, Christie Hospital NHS Foundation Trust, Manchester Academic Health Science Centre, University of Manchester, Manchester M20 4BX, UK.

[3]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Diseases, Botnar Research Centre, University of Oxford, Oxford OX3 7LD, UK.

[4]School of Cancer Sciences, University of Birmingham, Birmingham B15 2TT, UK.

[5]Genito Urinary Cancer Research Group, Division of Cancer Sciences, Faculty of Biology, Medicine & Health, University of Manchester, Manchester M20 4GJ, UK.

[6]Department of Urology, The Christie NHS Foundation Trust, Manchester M20 4BX, UK.

[7]Department of Histopathology, Leeds Teaching Hospitals NHS Trust, Leeds LS9 7TF, UK.

[8]Leeds Cancer Centre, St James's University Hospital, Leeds LS9 7TF, UK.

[9]Experimental Cancer Medicine Centre, Department of Oncology, University of Oxford, Oxford OX3 7DQ, UK.

[10]Department of Cellular Pathology, Manchester University Foundation Trust, Manchester M13 9WL, UK.

[11]Department of Cellular Pathology, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Oxford OX3 9DU, UK.

[12]NIHR Oxford Biomedical Research Centre, Oxford, OX3 9DU, UK.

[13]Weill Cornell Medical College, Cornell University, New York, NY, USA.

[14]Clinical Trials and Statistics Unit, Institute of Cancer Research, London SM2 5NG, UK.

[15]Academic Uro-Oncology Unit, The Royal Marsden NHS Foundation Trust, Sutton, London, SM2 5PT, UK; The Institute of Cancer Research, London, SM2 5NG, UK.

[16]Cancer Centre, Mount Vernon Hospital, Northwood, Middlesex HA6 2RN, UK.

[17]Manchester Cancer Research Centre, University of Manchester, Manchester M20 4GJ, UK.

Correspondence to:

Prof Anne E Kiltie
CRUK/MRC Oxford Institute for Radiation Oncology
Department of Oncology
University of Oxford
Old Road Campus Research Building
Off Roosevelt Drive
OXFORD OX3 7DQ

Tel. 01865 617352

anne.kiltie@oncology.ox.ac.uk

Tel. +44 1865 617352

Running title: MRE11 as biomarker in bladder cancer

Word count: [2606]

Conflict of Interest Notification

Dr Bhattarai, Dr Browning, Prof Choudhury, Prof Hall, Prof Huddart, Prof James, Prof Kiltie, Dr Strauss, Prof West report grant funding from Cancer Research UK during the conduct of the study.

Prof Choudhury is a Senior Editor of IJROBP and reports non-financial support from ASCO, non-financial support from ESTRO, personal fees and non-financial support from Bayer PLC, grants from Prostate Cancer UK, grants from Medical Research Council UK, grants from National Institute of Health Research UK, grants, personal fees and non-financial support from Elekta AB, personal fees from ASTRO, outside the submitted work; .

Prof Hall reports grants from Merck Sharp & Dohm, grants and non-financial support from Astra Zeneca, grants from Janssen-Cilag, grants and non-financial support from Bayer, grants from Aventis Pharma Limited (Sanofi), grants from Accuray Inc., outside the submitted work.

Prof Huddart reports grants and personal fees from MSD, grants and personal fees from Roche, personal fees from Bristol Myers Squibb, personal fees from Janssen, grants from Elekta, outside the submitted work.

Dr Iyer reports personal fees from Bayer Pharmaceuticals, outside the submitted work.

The authors have no other conflicts of interest to declare.

Funding acknowledgements

This work was funded by CRUK Biomarkers and Imaging Discovery and Development grant (C15140/A13492). The BC2001 trial was supported by grants (C547/A2606, C547/A6845, C9764/A9904, and C1491/A9895) from Cancer Research UK. The work was also supported by CRUK grant C2094/A11365, CRUK funding to the Cancer Research Manchester Centre (C147/A18083, C147/A25254), the NIHR Manchester, NIHR Oxford (Molecular Diagnostics Theme/Multimodal Pathology Theme) and Royal Marsden/Institute of Cancer Research Biomedical Research Centres. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

MRE11 as a predictive biomarker of outcome following radiotherapy in bladder cancer

Running title: MRE11 as biomarker in bladder cancer

ABSTRACT (206 words)

Purpose: Organ-confined muscle-invasive bladder cancer (MIBC) is treated with cystectomy or bladder preservation techniques, including radiotherapy. There are currently no biomarkers to inform management decisions and aid patient choice. Previously we showed high levels of MRE11 protein, assessed by immunohistochemistry (IHC), predicted outcome following radiotherapy but not cystectomy. Therefore, we sought to develop the MRE11 IHC assay for clinical use and define its relationship to clinical outcome in samples from two major clinical trials.

Methods and Materials: Samples from the BCON and BC2001 randomised controlled trials and a cystectomy cohort were stained using automated IHC methods and scored for MRE11 in three UK centres.

Results: Despite step-wise creation of scoring cards and standard operating procedures for staining and interpretation, there was poor inter-centre scoring agreement (Kappa 0.32, 95% CI 0.17-0.47). There were no significant associations between MRE11 scores and cause-specific survival (CSS) identified in BCON (n=132) and BC2001 (n=221) samples. Re-optimised staining improved agreement between scores from BCON tissue microarrays (n=116), but MRE11 expression was not prognostic for CSS.

Conclusions: Manual IHC scoring of MRE11 was not validated as a reproducible biomarker of radiation-based bladder preservation success. There is a need for automated quantitative methods and/or a reassessment of how DNA-damage response relates to clinical outcomes.

Keywords: predictive biomarkers, bladder cancer, immunohistochemistry, MRE11, NBN.

INTRODUCTION (2975 words)

Non-metastatic muscle-invasive bladder cancer (MIBC) can be treated with curative intent by either cystectomy or bladder preservation techniques, including radiotherapy alone or with a radiosensitising agent if tolerated (1, 2); neoadjuvant cisplatin-based chemotherapy is often used for fit patients. These approaches have a 40-60% cause-specific survival (CSS) rate at three years (3, 4). With no randomised data available, treatment is currently based on patient choice after discussion with a urologist, oncologist and nurse specialist (https://www.nice.org.uk/guidance/ng2). To date there are no validated biomarkers to predict the likely patient benefit from either approach (5).

We previously showed using immunohistochemistry (IHC) on pre-treatment transurethral resection of bladder tumour (TURBT) specimens that high levels of MRE11, a DNA-damage signalling protein, predicted outcome following radical radiotherapy for MIBC in two independent cohorts but not following cystectomy (6). Our results were subsequently independently validated in a Danish/German study (7).

The aims of the present study were to a) evaluate the ability to standardise the MRE11 immunohistochemistry (IHC) assay and its scoring methodology across multiple centres in the United Kingdom, thus developing it to appropriate standards for clinical use (final stage of biomarker discovery and assay development phase) (8) and, b) to again validate its correlation with outcome, in two of the largest and most important recent randomised trials of bladder preservation in MIBC (prospective analysis of retrospective tissue collections; first stage of biomarker qualification phase). REMARK guidelines were followed (8).

MATERIALS AND METHODS

Ethical approval was obtained from the National Research Ethics services in XXXX (project 09/H1013/24 and 10_NOCL_O1), XXXX (09/H0606/5) and XXXX (REC 00/8/75). All trial patients consented to use of their tissue and data for research.

Study populations. Patients in the UK multicentre randomised controlled trials BCON and BC2001 were given radiotherapy (RT) as 64 Gy in 32 fractions over 6.5 weeks or 55 Gy in 20 fractions over 20 weeks. BCON patients were randomised between RT alone or RT with carbogen and nicotinamide (samples available *N*=132 whole mount and 116 TMA samples) (1). BC2001 (CRUK/01/004) patients were randomised to RT alone or RT with 5-fluorouracil and mitomycin C (*N*=317 samples; split into test (n=154) and validation (n=163) cohorts) (2). One hundred samples were obtained from a cystectomy series at XXXX (9).

Materials. Pre-treatment formalin-fixed paraffin-embedded (FFPE) transurethral resection of the bladder tumour (TURBT) samples (see Table S1) were available for the trial and cystectomy cohorts. For BCON, whole FFPE blocks and 1 mm core tissue microarrays (TMAs) from invasive areas (constructed in Centre B in 2011) were used. Two in-house 'BIDD' TMAs containing 0.6 mm diameter cores were created for assay development in Centre A (Supplementary Material). Sections (4 μm thick) were cut and stored at 4^{o}C prior to use, with IHC performed no later than one month after cutting. A consultant uro-pathologist outlined areas of urothelial carcinoma invading the lamina propria (T1) and/or muscularis propria (T2) on haematoxylin and eosin- or MRE11-stained sections. Tumours with divergent differentiation within the invasive component were regarded as invasive urothelial carcinoma.

MRE11 immunohistochemistry. Following pilot work, a standardised operating procedure (SOP) was produced for MRE11 IHC using a Leica BOND-MAX™ autostainer (Leica Microsystems GmbH, Wetzlar, Germany) according to the manufacturer's instructions in Centre A, Centre B and Centre C (see Supplementary Materials). Slides were dewaxed in Bond Dewax solution (AR9222, Leica Microsystems) and rehydrated through graded ethanol

and distilled water. Tissue sections were washed using Bond Wash solution (AR9590, Leica Microsystems). Endogenous peroxidases were blocked using peroxidase block solution for 5 min, followed by antigen retrieval at pH 6 using Epitope Retrieval 1 solution (AR9961, Leica Microsystems) for 20 min at 100℃. Slides were then incubated with mouse monoclonal anti-MRE11 antibody (1:3,000, Abcam plc, Cambridge, UK, ab214, 1 mg/ml) for 15 min at room temperature. Primary antibody binding to tissue sections was visualised using a biotin-free Bond polymer refine detection system (DS9800, Leica Microsystems). After post-primary amplification for 8 min and detection with polymer for 8 min using 3,3'-diaminobenzidine for 10 min, slides were counterstained with haematoxylin for 1 min.

Six control slides were included in every 30-slide run. Negative control samples were stained using a mouse monoclonal immunoglobulin (Dako, Glostrup, Denmark, X0931, 100 mg/l) diluted to the same concentration as the MRE11 antibody and consisted of two patient samples from the cohort being stained and one slide from the BIDD TMA and a commercial sample. Positive controls consisted of sections from three of the commercial bladder tumour samples and a BIDD TMA section stained with the MRE11 antibody.

During the study, the automated IHC was improved by adding a 30 min 10% bovine serum albumin (BSA) pre-primary antibody protein blocking step, increasing the primary antibody dilution from 1:3,000 in 1% BSA to 1:6,000 in 10% BSA and reducing the primary antibody incubation time from 15 min to 8 min. Post re-optimisation, IHC was repeated on the BCON TMAs (final SOP in Supplementary Material).

In Centre D BCON samples were stained using an Autostainer Link 48 instrument (Dako, Inc) and an EnVision™ FLEX+, Mouse, High pH kit (Dako, K8002). Slides were deparaffinised and pre-treated in the automated Dako PT Link system using heated Envision Flex target high pH retrieval solution. Endogenous peroxidases were blocked using Flex Peroxidase Block for 5 min. Slides were incubated with mouse monoclonal anti-MRE11 antibody (1:3,000, Abcam plc, ab214, 1mg/ml) for 30 min at room temperature. After

applying labelled polymer Flex/horseradish peroxidase for 20 min, the staining was visualised using Flex DAB+ substrate chromogen for 2x5 min, and slides were counterstained with haematoxylin.

MRE11 scoring. Slides were scanned using an Aperio ScanScope CS scanner (Leica Microsystems) at 40x magnification and visualised by ImageScope Viewer. Scoring similar to that described by us elsewhere (6) was undertaken following training, with guidance sought on IHC interpretation from a histopathologist for challenging cases (details in Supplementary Material). Briefly, six to 10 images (containing at least 100 cells) were taken from random fields within the invasive areas. Surface papillary tumours and carcinoma in situ were not scored. Care was taken to avoid taking images from areas distorted/damaged by diathermy or crush artefact and from necrotic areas, reducing the potential to include cells with unreliable immunostaining. Tumour nuclear MRE11 staining intensity was graded 0 to 3+ (Figure S1a) by 2-3 independent scorers within each centre using a guide comprising 30 images. The modal intensity for each of the 6-10 images was determined and an overall modal intensity score assigned to each case. Comparison was then made of results from individual scorers, differences highlighted and a consensus reached for each centre. Percentage positivity was determined by either (initially) counting 100 cells per image using ImageJ software (Bethesda, MD) (10) or using a second standardisation scoring guide with 30 images to estimate the percentage positivity. The mean percentage of positive cells was multiplied by the modal intensity to give a semiquantitative H-score (0-300, Table S2).

Statistical Analysis. Analyses were conducted independently via STATA in each cohort, with a 25% cut-off used to allow comparison with previous publications. Within each centre, for a cohort**,** the inter-rater agreement of MRE11 intensity scores was assessed using the weighted Kappa statistic via STATA kappaetc packages. The reliability of percentage positive scores was studied using intraclass correlation coefficient (ICC). A higher value indicates a better agreement between scorers (11). Between-centre, the reliability of MRE11 intensity and percentage positive scores were assessed by the Kappa statistics and the ICC,

respectively. Associations between MRE11 H-score (≤25th percentile and >25th percentile) and bladder cancer specific survival (CSS) were analysed using Kaplan-Meier graphs with log-rank tests. In the BCON TMA cohort, a subgroup analysis was conducted for radiotherapy patients only. Hazard ratios (HR) were generated for MRE11 H-score >25th percentile using Cox regression with adjustment for treatment, stage, grade, completion of TURBT, pre-treatment haemoglobin level and number of radiotherapy fractions.

RESULTS

A schematic of the study design is presented in Figure 1.

Staining and scoring reliability. The MRE11 assay was developed in Centre A (see original Standard Operating Procedure). Staining was highly reproducible between runs. A working dilution of 1:3,000 was agreed by aa, bb and ff. A range card was created from the MIBC samples from Centre A to represent 0, 1+, 2+ and 3+ intensity scores for tumour nuclear MRE11 expression (Figure S1). A 30-sample PowerPoint slide deck was then developed for subsequent scoring by aa, bb, gg and hh (the latter two were experienced in manual scoring [Ref 6]). Blinded scoring resulted in concordance between scores from four observers in 57% of the 30 cases (87% agreement for three scorers). Following discussion, agreement was achieved for all samples.

Initially 40 BCON slides (parallel whole sections) were stained and scored in Centre B and Centre A and 7 months later a further 132 slides were stained and scored. Staining was similar between centres, antibody aliquots (Table 1 and Figure S2), and across time (Figure 2a). However, scores were higher in Centre A with 87 of the 132 samples (66%) scored as 3+ versus 50 (38%) in Centre B (Table 2a). The data equated to H-score 25th centile cut-points of 175.8 for Centre B and 204.1 for Centre A; in the test and validation cohorts of XXXX [Ref 6], these had been 130 and 76 respectively.

We concluded that automated staining and improved imaging resulted in higher scores than the manual methods previously used. The 1+/2+ and 2+/3+ cut-point boundaries were then redefined and a revised 30-slide template produced. Rescoring reduced the number of 3+ scores to 42/132 (32%, Centre A) and 17/132 (13%, Centre B) (Table 2a) and respective median H-scores from 197 to 148 and 195 to 149 (Table S2).

Before the data were linked with survival outcomes, a more formal assessment of standardisation of scoring methods was undertaken. Twenty randomly selected Centre A images were scored by two or three Centre B observers (xx, yy, gg), consensus reached

and comparison made with Centre A scores; a similar procedure was performed for 20 randomly selected Centre B images scored by three Centre A observers (aa, bb, cc). Centre B scores tended to be lower than Centre A. Three of 20 Centre B and three of 20 Centre A cases were discordant, although consensus was reached in all but one of the latter (Table S3).

One hundred and fifty-four BC2001 test slides were stained and 144 scored in both Centre A and Centre C. The Centre C observer scored 102 (66%) samples as 3+ compared to 44 (29%) in Centre A. In Centre A, a further 163 BC2001 validation slides were stained and imaged (by bb) and staining intensity and estimated positive percentage were scored in 145 slides independently by 2 observers (aa and bb) and consensus scores reached.

Associations with cancer-specific survival. Analyses carried out using data generated with the initial staining procedure in BCON samples showed no significant associations between MRE11 expression and CSS (Figure S4a and b). Analysis of evaluable samples from 221 training and validation BC2001 patients by Centre A confirmed the lack of significant association between MRE11 expression and CSS (log rank test P=0.97, Figure S4c). Analysis of 99 patients in the cystectomy cohort showed a lack of significant association between MRE11 expression and CSS (log rank test P=0.19, Figure S4d).

Refinement of the assay. A subset of BCON samples (n=48) stained in the Centre D Pathology Department (Figures 2b and c and S3) showed reduced non-specific background staining compared to samples stained in Centre A and Centre B (Table S4). Additionally, in 20 of these samples there was an increase in the variation of staining intensities between tumours. Therefore, at Centre A attempts were made to improve the automated IHC staining on the BOND instrument to replicate that on the Dako processor. This involved adding a pre-primary antibody protein blocking step, increasing the primary antibody dilution to 1:6,000 and reducing the primary antibody incubation time (see Methods for details, final SOP in Supplementary Material). Following re-optimisation, staining was repeated on 116 TMA

samples from BCON (Table 3). The re-optimised staining protocol resulted in moderate to high agreement between scorers (Figure 2c, Table 1).

In the 116 BCON TMA patients, there was no significant association between MRE11 expression and CSS (Figure 3a and b). We hypothesised that the use of carbogen and nicotinamide could have influenced the response to radiotherapy via hypoxia modification. Therefore, a subgroup analysis of the 62 patients who received radiotherapy was performed. This analysis displayed a non-significant trend for an association with CSS when expression was scored in Centre A (Log rank test, P=0.20) but not seen in scores from Centre B (Figure 3c and d).

DISCUSSION

This study aimed to develop the MRE11 IHC assay for prospective clinical use. It was hoped the work would underpin the development of a trial randomising patients to conventional versus MRE11-guided patient choice, and subsequent introduction of routine MRE11 testing in the NHS. IHC is an attractive platform for clinical use as illustrated by HER2 testing (12), but routine implementation requires rigorous validation of the staining and scoring methods to ensure consistency and reliability across institutions (13). Automated staining aids standardisation and efficiency by improving fidelity and workflow (14).

We present the first attempt to validate MRE11 IHC staining across centres using good clinical laboratory practice standards. This was a large collaborative effort and we provide a robust level of validation. The findings highlight the challenges associated with standardising an MRE11 IHC test. Staining was qualitatively reproducible between centres, but scoring was not. Although histopathologists did not score the samples, they provided training, input on interpretation and arbitration on challenging cases. Problems with inter-observer scoring agreement were highlighted by the Ki67 Working Group (15) where scoring in 22 laboratories in 11 countries yielded ICC values ranging from 0.84 to 0.93. Our ICC values for density scores were similar, ranging from 0.90 to 0.98, but there were discrepancies in intensity scores despite efforts to improve agreement between laboratories. Potential reasons for the poor concordance include insufficient training, potential subjective bias or technical factors such as differences in screen resolutions.

Scores generated by a single scorer in Centre C (using the scoring cards generated but with less intensive training) were more discordant than those generated in Centre A and Centre B, reflecting the need for external quality assessment schemes. Whilst training can improve levels of concordance, e.g. for EGFR staining (16), some stains are intrinsically more difficult to score than others (17), including MRE11. Another issue is time taken for scoring, with the Centre D Consultant Histopathologist taking 25-30 minutes per case (see Supplementary Material). Automated digital image analysis might remove human scoring bias and allow

rapid scoring of multiple samples (18). We attempted to score a subset of samples using automated digital analysis, but found that when optimised to higher intensity samples, the results for low intensity samples were inaccurate, and vice versa. Therefore, we did not pursue this. It may be that the development of more sophisticated algorithms might obviate this problem.

We failed to validate the previous findings from us and others (6, 7). Although this could be due to lack of biological effect, we think it more likely to be due to methodological issues, including problems in standardising the automated staining and poor scoring reproducibility across centres due to difficulties in standardising intensity scoring. Others found similar difficulties when studying ERCC1 expression using an 8F1 antibody in a large sample set from two phase III trials of adjuvant cisplatin in lung cancer. A change in the batch of antibody used resulted in an inability to validate the predictive effect of ERCC1 immunostaining (19). From a biological point of view, it appears paradoxical that high expression of a DNA damage signalling protein (MRE11) might be associated with better outcomes following a DNA damaging agent (ionising radiation). However, we recently observed a truncated version of MRE11 in a bladder cancer cell line which is still detected by the antibody used in this study (20), and we hypothesise that this might act in a dominant-negative fashion. This hypothesis is currently under investigation in a separate study.

With the re-optimised staining method, due to sample depletion, we only obtained MRE11 data on 116 BCON patients, which provided only 82% power to detect a change in 3-year CSS between 43% and 70% between two MRE11 groups, as reported previously (6). A non-significant trend for a difference in CSS was seen in patients receiving radiotherapy alone, but only in the Centre A stained and scored samples. Carbogen and nicotinamide are given to reduce hypoxia within tumours and, by increasing the biological effectiveness of radiotherapy (21, 22), could modify the association between MRE11 and CSS. With small numbers of patients, the radiotherapy alone subgroup analysis was underpowered.

Despite our failure to validate MRE11 as a prognostic marker in radiotherapy patients, we cannot reject a role for MRE11 as a biomarker in MIBC. Indeed, a meta-analysis of the BCON radiotherapy alone with data obtained from 44 bladder chemoradiation patients from Centre E (see Supplementary Materials for methods and results, Fig S5) yielded a pooled hazard ratio for MRE11>25th percentile of 0.47 (95% CI 0.13 to 1.03). This HR is similar to those reported previously (0.42) (6) and 0.64 (7). Furthermore, recently the RTOG group has taken an alternative approach to scoring MRE11 using an internal control of the nuclear:cytoplasmic ratio of MRE11 and the more standardisable automated quantitative analysis (AQUA) approach (23) with promising results. Indeed, other DNA damage response genes have proved more tractable. For example, a recent microscopy-based nucleotide excision repair assay to profile *ERCC2* mutations established a role for *ERCC2* helicase domain mutations as a predictive biomarker in bladder cancer treated with cisplatin-based chemotherapy. *ERCC2* mutational status has now been incorporated as a predictive biomarker in risk-adapted MIBC clinical trials (24).

Limitations of our study include the eventual reduced statistical power due to sample attrition and possibly use of TMAs. However, studies have identified concordance between IHC scoring of ≥0.6mm TMAs and whole tissue sections, especially with multiple same-patient cores (25, 26). It is therefore reasonable to assume comparable results between the BCON whole tissue sections and 1 mm TMAs used here. In our study, death from other causes was treated as censored in the analysis. We attempted to apply competing risk analyses that take into account death from other causes. However, these resulted in similar findings to the analyses presented and only 4/44 people in the Centre E cohort died from other causes. Therefore, we chose not to formally present competing risk analyses.

Numerous studies have identified potential IHC-based biomarkers but only a few have obtained FDA approval (13). Despite the reduced stability of RNA versus protein, it is easier to measure at low abundance and with greater sensitivity and specificity (27). Generation of

a gene signature that reflects MRE11 protein expression might provide a more robust biomarker than IHC. Further exploration of a gene classifier would be worthwhile.

CONCLUSIONS

In this study, we were unable to validate MRE11 as a robust, reproducible, predictive biomarker for radiotherapy response in MIBC. A large analysis of prospectively acquired tissue is required using the refined staining methodology along with further exploration of automated digital scoring methods. Alternatively, biomarkers based on other proteins or genomic data may be better placed for clinical use in future.

REFERENCES

1. Hoskin PJ, Rojas AM, Bentzen SM, Saunders MI. Radiotherapy with concurrent carbogen and nicotinamide in bladder carcinoma. J Clin Oncol. 2010;28(33):4912-8.

2. James ND, Hussain SA, Hall E, et al. Radiotherapy with or without chemotherapy in muscle-invasive bladder cancer. N Engl J Med. 2012;366(16):1477-88.

3. xxx

4. Booth CM, Siemens DR, Li G, et al. Curative therapy for bladder cancer in routine clinical practice: a population-based outcomes study. Clin Oncol (R Coll Radiol). 2014;26(8):506-14.

5. Forker LJ, Choudhury A, Kiltie AE. Biomarkers of Tumour Radiosensitivity and Predicting Benefit from Radiotherapy. Clin Oncol (R Coll Radiol). 2015;27(10):561-9.

6. xxx

7. xxx

8. McShane LM, Altman DG, Sauerbrei W, et al. REporting recommendations for tumour MARKer prognostic studies (REMARK). Brit J Cancer. 2005;93(4):387-91.

9. Ramani VAC, Maddineni SB, Grey BR, Clarke NW. Differential Complication Rates Following Radical Cystectomy in the Irradiated and Nonirradiated Pelvis. Eur Urol. 2010;57(6):1058-63.

10. Abramoff MD MP, Ram SJ. Image processing with Image J. Biophotonics International. 2004;11(7):36-42.

11. Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. Tutor Quant Methods Psychol. 2012;8(1):23-34.

12. Perez EA, Cortes J, Gonzalez-Angulo AM, Bartlett JM. HER2 testing: current status and future directions. Cancer Treat Rev. 2014;40(2):276-84.

13. O'Hurley G, Sjostedt E, Rahman A, et al. Garbage in, garbage out: a critical evaluation of strategies used for validation of immunohistochemical biomarkers. Mol Oncol. 2014;8(4):783-98.

14. Howat WJ, Lewis A, Jones P, et al. Antibody validation of immunohistochemistry for biomarker discovery: recommendations of a consortium of academic and pharmaceutical based histopathology researchers. Methods. 2014;70(1):34-8.

15. Leung SCY, Nielsen TO, Zabaglo L, et al. Analytical validation of a standardized scoring protocol for Ki67: phase 3 of an international multicenter collaboration. NPJ Breast Cancer. 2016;2:16014.

16. Ruschoff J, Kerr KM, Grote HJ, et al. Reproducibility of immunohistochemical scoring for epidermal growth factor receptor expression in non-small cell lung cancer: round robin test. Arch Pathol Lab Med. 2013;137(9):1255-61.

17. xxx

18. Khosravi P, Kazemi E, Imielinski M, et al. Deep Convolutional Neural Networks Enable Discrimination of Heterogeneous Digital Pathology Images. EBioMedicine. 2018;27:317-28.

19. Friboulet L, Olaussen KA, Pignon JP, et al. ERCC1 isoform expression and DNA repair in non-small-cell lung cancer. N Engl J Med. 2013;368(12):1101-10.

20. xxx

21. Hoskin PJ, Saunders MI, Phillips H, et al. Carbogen and nicotinamide in the treatment of bladder cancer with radical radiotherapy. Br J Cancer. 1997;76(2):260-3.

22. Laurence VM, Ward R, Dennis IF, Bleehen NM. Carbogen breathing with nicotinamide improves the oxygen status of tumours in patients. Br J Cancer. 1995;72(1):198-205.

23. Magliocco AM, Moughan J, Simko J, et al. The impact of MRE11 in nuclear to cytoplasmic ratio on outcomes in muscle invasive bladder cancer an analysis of NRG/RTOG 8802, 8903, 9506, 9706, 9906, and 0233. Journal of Clinical Oncology. 2017;35(6_suppl):343-.

24. Li Q, Damish AW, Frazier Z, et al. ERCC2 Helicase Domain Mutations Confer Nucleotide Excision Repair Deficiency and Drive Cisplatin Sensitivity in Muscle-Invasive Bladder Cancer. Clin Cancer Res. 2019;25(3):977-88.

25. Camp RL, Neumeister V, Rimm DL. A decade of tissue microarrays: progress in the discovery and validation of cancer biomarkers. J Clin Oncol. 2008;26(34):5630-7.

26. Voduc D, Kenney C, Nielsen TO. Tissue microarrays in clinical oncology. Semin Radiat Oncol. 2008;18(2):89-97.

27. Xi X, Li T, Huang Y, et al. RNA Biomarkers: Frontier of Precision Medicine for Cancer. Noncoding RNA. 2017;3(1).

FIGURE LEGENDS

Figure 1: Schematic of experiments undertaken.

Figure 2: **Comparison of MRE11 staining** between a) initial (May) Centre A and subsequent (December) Centre A and Centre B staining of BCON samples; b) Centre D, Centre A and Centre B staining of BCON samples; c) re-optimised staining (Centre A new) on BIDD TMA compared to Centre D staining and original Centre A staining (Centre A old)

Figure 3: Kaplan-Meier survival plots for MRE11 expression >25th centile or =< 25th centile for a) Centre A whole cohort; b) Centre B whole cohort; c) Centre A radiotherapy alone subgroup; d) Centre B radiotherapy alone subgroup BCON TMAs.

Table 1: Intra- and inter-centre scoring agreement. * % positive cells were scored by one scorer; ** Not possible to calculate as there was very little variation in slides for one scorer.

Comparison	**Initial staining**		**New staining**	
	Kappa (95% CI) on intensity	ICC (95% CI) on % positive cells	Kappa (95% CI) on intensity	ICC (95% CI) on % positive cells
BCON (n=132)				
Intra-centre: Centre A scorers	0.63 (0.52, 0.73)	NA*	0.59 (0.49, 0.68)	0.90 (0.88, 0.93)
Intra-centre: Centre B scorers	0.75(0.64, 0.85)	NA *	0.95 (0.91, 0.99)	0.98 (0.98, 0.99)
Inter-centre : Centre A –Centre B	0.32 (0.17, 0.47)	0.73 (0.64, 0.82)	0.55 (0.41, 0.70)	0.91 (0.88, 0.94)
BC2001				
Testing (n=144)	0.66 (0.58, 0.74)	0.95 (0.93,0.96)	NA	NA
Validation (n=145)	0.72 (0.62, 0.83)	NA**	NA	NA

Table 2: Comparison of intensity scores for BCON and BC2001 cohorts. a) BCON using original scoring cutpoints Centre A vs Centre B and using revised scoring cutpoints Centre A vs Centre B; b) BC2001 test cohort Centre A vs Centre C (zz) and BC2001 validation cohort Centre A).

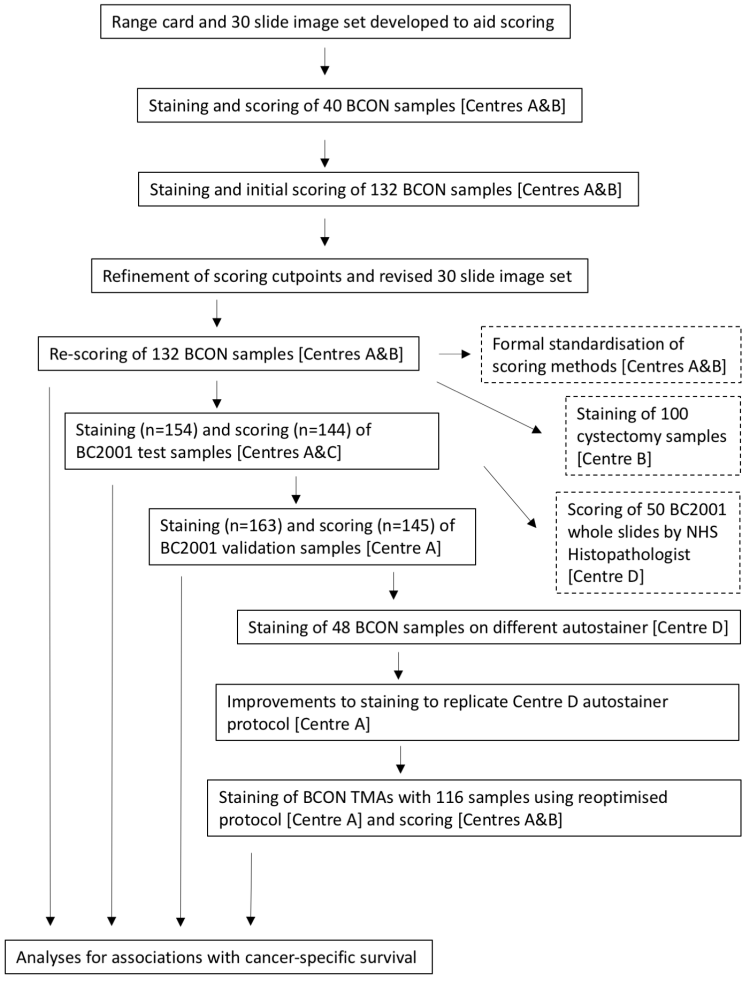
BCON original			BCON revised	
intensity	Centre A	Centre B	Centre A	Centre B
0	0	0	0	0
1+	4	18	11	22
2+	24	48	62	77
3+	87	50	42	17
not good	17	16	17	16
total	132	132	132	132

b)

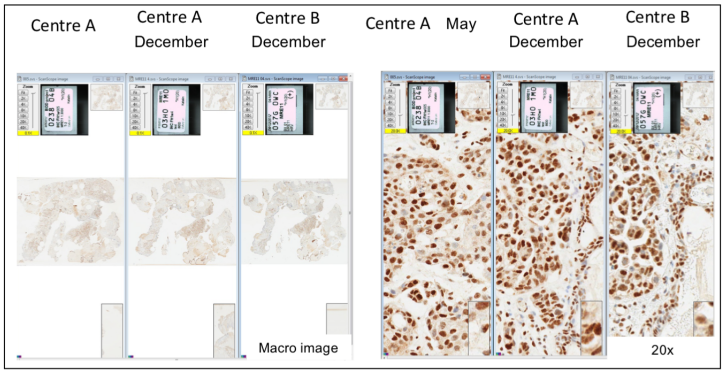
BC2001 test cohort			BC2001 validation cohort
intensity	Centre A	Centre C	Centre A
0	0	0	0
1+	19	3	8
2+	81	33	81
3+	44	102	56
not good	10	16	18
total	154	154	163

Table 3: Baseline characteristics of the 116 patients in the BCON TMA dataset.

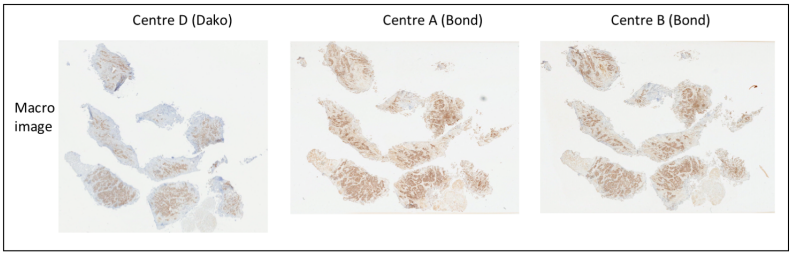
Characteristics	**Number (%) / Median (Range)**
Median age (range)	75.65 (51.5 to 90.5)
Gender	
Male	101 (87.07%)
Female	15 (12.93%)
Tumour Stage	
T1	5 (4.31%)
T2	83 (71.55%)
T3	23 (19.83%)
T4a	5 (4.31%)
Tumour:stromal ratio(TSR)	
High	109 (96.46%)
Low	4 (3.54%)
Growth margins	
Broad	4 (3.54%)
Infiltrative	109 (96.46%)
Growth patterns	
Both	47 (41.59%)
Papillary	10 (8.85%)
Solid	56 (49.56%)
Necrosis	
No	47 (41.59%)
Yes	66 (58.41%)



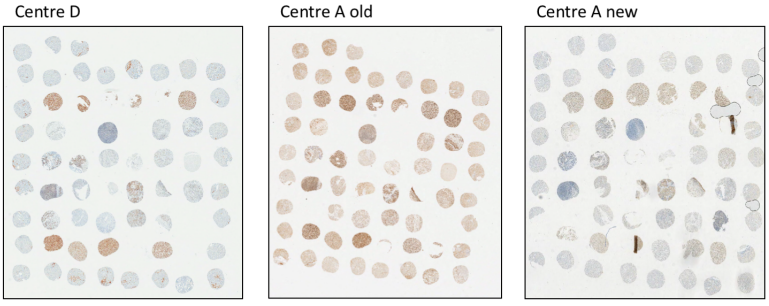
a)

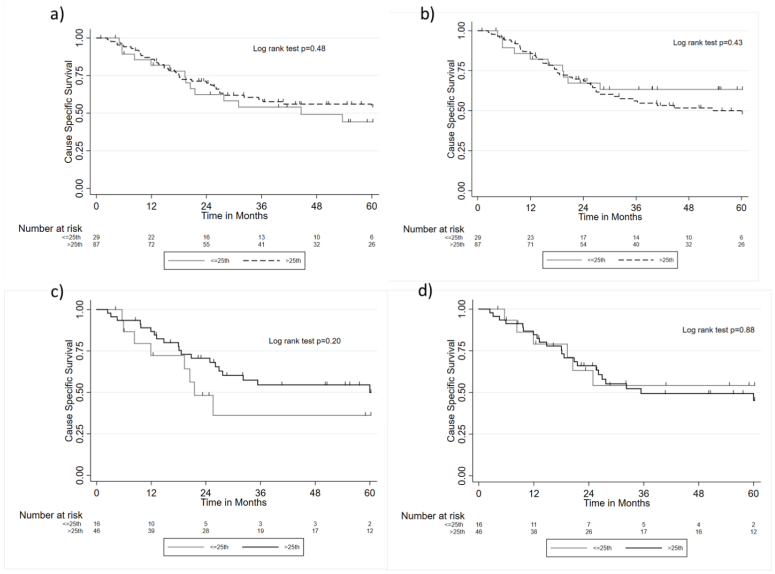


b)



c)





SUMMARY

We attempted to develop an MRE11 IHC assay to appropriate standards for clinical use and to define the relationship of the biomarker to clinical outcome in a prospective analysis of retrospective tissue from two randomised trials. Staining was reproducible across centres with scoring less so, so we could not validate MRE11 as a robust, reproducible predictive biomarker of radiotherapy response. Our study demonstrates the challenges involved in developing a robust IHC-based protein biomarker.