Online-only material: Supplemental Methods

Patient preparation
On the day before surgery, fasting patients underwent placement of an urinary bladder catheter immediately prior to the positron emission tomography/magnetic resonance imaging (PET/MRI) examination to avoid size differences of the urinary bladder in PET and MRI scan and to reduce bladder \([^{18}F]\)fluoroethylcholine chloride (FEC)-activity overlay of the prostate. Urological data, such as prostate specific antigen (PSA) concentration, and the results of digital rectal examination, transrectal sonography and needle biopsy were documented by an experienced urologist in the urological case report form. After applying the endorectal MRI coil, patients were positioned on a vacuum mattress on the MRI table to minimize the mismatch in the subsequent software fusion procedure. Additionally, four PET/MRI multimodality fiducial spot markers containing 37kBq \([^{22}Na]\) and a MRI T2 weighted (T2w) hyperintense gel (Eckert & Ziegler Isotope Products, Inc., Valencia) were attached at the hip region to facilitate landmark PET/MRI fusion.

Magnetic Resonance Imaging
The MRI examination was performed using a 1.5-Tesla MRI system (Gyroscan ACS-NT, Philips, Hamburg, Germany) with combined QBody and endorectal coil. Pelvic assessment and lymph node staging was effected with 5 mm T2w turbo spin echo (TSE) transversal and coronal short-tau inversion recovery (STIR) sequences. For prostate assessment, 3 mm endorectal T2w spin echo (SE) sagittal, transversal, and coronal sequences were acquired. Transversal sequences were angulated 90° to the intraprostatic bladder catheter in order to allow exact correlation with histological holoptical slices. A reference slice was determined in the center part of the organ, measured from the prostate base. This data served for the pathologist as a standard of reference for the first cut of the resection preparation. Images were assessed by consensus of two experienced MRI-radiologists (WB, 12 years; BD, 16 years), who were blinded to the PET images and patient’s urological data, but aware of the study inclusion criteria. MRI-only analysis included the diagnosis of suspected tumor, T2w-hypointense intraprostatic lesions, local staging of tumor extent, and lymph node staging of the pelvic region. The assessment of the peripheral zone as well as the transitional and periurethral zone was according to criterias later composing in part the Pi-RADS(1) score, which did not yet exist during the data collection for this study. Hypointense, unsharp lesions anywhere in the prostate were assessed as malignant (todays Pi-RADS IV and V). Indistinct or ambiguous areas (todays Pi-RADS III) required a consensus decision between the two radiologists insofar as the case report form had only the binary choice ‘benign changes’ or ‘malignant focus’ for each axial slice. Tapered, hypointense lesions of the peripheral zone (todays Pi-RADS II) were assessed as benign, but hypointense, sharp-edged lesions of the central gland also required a consensus decision of the two radiologists. Extraprostatic extension (EPE) of local tumor foci was suspected according to direct and indirect MRI signs if the prostate margins were blurred or pre-bulged, or when an extraprostatic hyposign was present in the surrounding tissue. Invasion of seminal vesicles (T3b) was suspected in cases of hypointense lesions of the prostate base continuing into the proximal parts or the opening of the ejaculatory duct of the seminal vesicles. Results were documented in the radiologic study case report form directly after image analysis. After MRI acquisition, the modular MRI table was released from the scanner, and lifted and fixed on the PET table system. The endorectal coil was not removed until the end of the PET investigation.

Rationale for the applied Radiopharmaceutical (FEC)(2)
Metabolic imaging of prostate cancer is challenging due to its heterogeneous histological behavior. To date, three radiotracers for positron emission tomography (PET) with different
pathophysiological behavior have been most studied in Nuclear Medicine, and have demonstrated varying utility for detecting this tumor entity: $^{18}$F-FDG, $^{18}$F- and $^{11}$C-acetate, $^{18}$F-(ethyl) and $^{11}$C-choline. Elevated uptake of $^{18}$F-FDG is based upon increased expression of the GLUT 1 transporter of the cellular membrane as well as enhanced hexokinase II enzymatic activity in tumors. Whereas $^{18}$F-FDG is the most common PET-radiotracer for general tumor imaging, its utility for imaging prostate carcinoma seems to be limited to the poorly differentiated subtypes. Arising from the suggestion that fatty acid metabolism rather than glycolysis is upregulated in prostate carcinomas, imaging with $^{18}$F- and $^{11}$C-acetate has been developed. Radiolabelled acetate is incorporated into phosphatidylcholine and neutral lipids of the cell and seems to have higher uptake in prostate carcinomas than does $^{18}$F-FDG. However, there also seems to be an age-related physiological uptake in normal prostate tissue as well. $^{18}$F-(ethyl)- and $^{11}$C-choline accumulate in prostate tumors due to an upregulation of choline kinase, which leads to incorporation and trapping in the form of phosphatidylcholine (lecithin). Both radiopharmaceuticals (acetate and choline) have demonstrated their utility for imaging prostate cancer in several retrospective studies, whereas more research had been done with radiolabelled choline. Therefore we chose this class radiotracer for a prospective clinical trial. Because $^{11}$C-choline has a very short physical half-life of only 20 minutes and therefore requires an on-site cyclotron/radioehmistry facility, we opted to use $^{18}$F-ethylcholine (FEC), which has a half-life of 110 minutes, making it a more practical radiotracer for the clinical routine.

**Positron Emission Tomography**
PET scans were performed with an LSO full-ring scanner (ECAT ACCEL, Siemens, Erlangen) using a multiphase protocol starting with a “cold” transmission scan of the lower pelvis. This was followed by a list mode emission scan with 10 frames of one minute each, starting immediately after the administration of 3.3 MBq per kilogram body weight of FEC (Eckert & Ziegler EURO-PET Berlin GmbH, Berlin, Germany) as a bolus through the cubital vein. After a short gap necessitated by computer processing time the whole body scan was performed starting at the upper thoracic aperture down to the proximal femur. Acquisition parameters were 3 minutes per emission scan and 2 minutes per transmission scan for each of the five bed positions. As a result, the prostate region was scanned again at $t_1 = 48 \pm 8$ minutes after tracer injection. A delayed local acquisition at $t_2 = 71 \pm 9$ minutes of the lower pelvis with a 6-minute emission and 2-minute transmission scan concluded the diagnostic acquisition procedure. Image reconstruction parameters consisted of two iterations and eight subsets for whole body and delayed scans and an additional four iterations and 16 subsets in delayed local scans for sharpening focal intraprosthetic choline uptake for improved visual analysis. Images were reconstructed in transversal, sagittal, and coronal planes. Image assessment was undertaken by consensus of two PET-experienced Nuclear Medicine physicians (MH, 11 years; BK, 13 years) who were aware of the inclusion criteria but blinded for the MRI and histological results. Intraprostatic focal uptake above the surrounding prosthetic background uptake was interpreted as malignant in PET-only analysis and semiquantitative analysis was documented as a standardized uptake value at maximum (SUV$_{max}$) and mean (SUV$_{mean}$) level (50% isoconture) using a dedicated software package (Siemens Syngo, Siemens Medical, Erlangen, Germany).

**Combined Positron Emission Tomography / Magnetic Resonance Imaging**
PET images at both time points were fused with transversal endorectal and QBody T2w MRI images using a dedicated landmark fusion tool (Multi Modality, Hermes Medical Solutions, Stockholm, Sweden), where the four PET/MRI fiducial spot markers and the arterial iliac vessels on early post-injection list mode PET images served as references. No significant registration mismatch occurred between both modalities due to above mentioned patient
preparation. Because the assessment of PET/MRI is a new approach, a strict a priori decision-making process on how to interpret the combined images was avoided in the initial study plan. Consequently, the following algorithm was prospectively applied as a consensus of all four readers (Nuclear Medicine and Radiology) and documented in the FEC-PET/MRI case report form immediately after the images were fused and before the patient underwent surgery. For PET/MRI analysis, MRI-suspect lesions (as defined by the criteria above) without FEC uptake were considered non-malignant. PET-positive and MRI negative lesions in the central periurethral zone were also considered to be benign, e.g. hyperplasia. Otherwise, MRI-suspect lesions in the central or peripheral zone with FEC uptake were assessed as malignant and FEC negative areas without suspect MRI hyposignal were considered benign. PET/MRI analysis was carried out according to these four rules. As far as these rules governed this analysis, the PET/MRI consensus of all four readers was mainly required for the correct alignment of the lesions assessed in the individual method’s CRFs. MRI-only and PET-only analysis and PET/MRI results including the PET-SUVmax/mean were documented in the nuclear medicine case report form before surgery and thus remained blinded to the histological results. The PET/MRI findings were prospectively printed as a hard copy for image documentation and added to the CRF facilitating e.g. the post-hoc ROC-analysis of the patient’s main tumor focus as presented.

**Histology**

After radical prostatovesiculectomy with pelvic lymphadenectomy by the urologist (CS, 31 years experience), a segment of the intraprostatic bladder catheter was kept in position within the excised surgical specimen during pressurized thermal fixation (four to six hours at 40°C with 4% formaldehyde) to allow an exact 90° cut of organ slices. The organ was rendered after fixation with different colors for each side. The reference plane from MR images was measured and taken as first cut. Transversal slice thickness was approximately 3-5 mm from apex to base. The first apex slice and the last base slice were additionally cut vertically for assessment of the apical or basal pT3a stage and the invasion of seminal vesicles as well as microscopic invasion of the bladder neck and the sphincter muscles (which did not occur in our cohort). After embedding and staining with hematoxylin-eosin, the prostate, seminal vesicles, and lymph nodes were assessed in consensus by two blinded experienced pathologists (KK, 32 years of experience; SH, four years). Focal extraprostatic extent (focal EPE) was assessed according to the criterias of Wheeler et al.(3): less than one high-power field on no more than two separate sections. Tumor edges were dot-marked and step sections were holoptically digitalized using an electronic microscope system (Leica Microsystems GmbH, Wetzlar, Germany) controlled by the Hilgers™ DISKUS software package, and results were documented in the pathological study case report form. For the re-assessment of the PET and MRI false positive findings, pathologists were unblinded after the case report form for pathology was closed. The benign findings were documented separately.

**Lesion alignment PET/MRI-Histology**

Lesions were assessed in a side-by-side analysis using a portable software fusion system (Hermes Medical Solutions, Stockholm, Sweden) next to an electronic microscope system (MH and SH in consensus) after identical angulation of the histological step sections and the MRI slices, and by using the reference slice for vertical allocation (as described above). A matching tumor lesion had to be in the same three-dimensional organ region in both imaging (PET, MRI, PET/MRI) and histology. In addition, the lesion shape as visualized on imaging had to fit the histological one (as demonstrated in Figure 1.a and 1.b). A tumor lesion was only counted once in histology (and declared as one larger tumor focus) if it was located in the same organ region the section above and/or below (e.g. as shown in the three-dimensional MPR PET/MRI reconstruction in Figure 1.a and 1.b). If two lesions in one section converged...
in the section above and/or below (axial cut of carcinoma branches) they were also declared as a single lesion. A tumor microfocus < 5 mm in diameter was defined as one lesion when present in only one section < 5 mm in diameter and not detected in either the section above or below. MRI-only, PET-only, and combined PET/MRI assessments were correlated with the histologically dot-marked tumor lesions. In cases of mismatch (no tumor lesion dot-marked in areas of PET and/or MRI positive imaging results), correlating areas in histology were re-assessed to classify the benign changes (consensus of KK and SH). A prospective, blinded assessment of all intraprostatic benign changes in histology would not have led to a valid analysis because these areas are ubiquitous, versatile and are rarely sharp-edged. Thus, the analysis of these benign areas was conducted by an additional assessment of all initial PET and MRI false positive lesions. Concerning MRI, this analysis was conducted in addition to the study protocol, where originally only PET-false positive imaging results were re-assessed. Taking into account that the spatial resolution of the PET scanner used in the present study was limited to 6 mm, tumor lesions > 5 mm in diameter were additionally assessed in a separate analysis.

**Statistical analysis**

For the primary endpoint of this study the number of patients was estimated on the following basis: The anticipated patient-based sensitivity in tumor detection of PET/MRI was ~90%; the desired diagnostic accuracy in determining this was at least 9%, so that the lower limit of the 95% confidence interval would be above 80%. The relationship between the lowest number of observed cases and the 95% confidence interval of a probability is given by \( N = \frac{1.96^2 \times p (1 - p)}{a^2} = 43 \), where \( p = 0.90 \) (the anticipated probability) and \( a = 0.09 \) (the desired accuracy for determining the probability) [Sachs, 1993]. Since the study was conducted with \( N = 38 \) patients, calculation of the formula above yielded \( a = 0.095 \) and in a lower limit of the 95% confidence interval (based on \( p = 0.9 \)) of 0.805. Therefore, the total of 38 patients as enrolled in the study was still sufficient to achieve a lower limit of the 95% confidence interval above 80%. Deviation from the original plan to include 43 patients occurred, as the sponsor decided to terminate the ongoing study when the study site received a new PET/CT scanner; it was not considered acceptable to continue using the old scanner merely for the sake of achieving the planned recruitment figure. Furthermore, inclusion of patients’ imaging results obtained on the new scanner would have impacted the homogeneity of the data set and thus imperiled the interpretability of the study. At that point, 38 patients had been recruited and treated, corresponding to 88% of the 43 patients planned. Based upon the power calculations above, this was considered an acceptable figure, and the study was therefore terminated at that point. The loss in information value arising from the analysis of 38 instead of 43 patients is not considered to have had a serious impact on the aims or the validity of the study.

Data are presented by means of summary and/or frequency tables. Summary tables present the number of observations as well as mean, standard deviation, minimum, median, and maximum, while frequency tables present absolute and relative frequencies. For determination of sensitivity, specificity, and accuracy, ‘2x2’ tables were used. These were checked for independence using Fisher’s exact test at a descriptive level. The respective 95% confidence intervals were calculated. When not at patient level, by taking into account the fact that the observational units (lesions) are 'clustered' within patients, the confidence intervals were calculated for clustered samples, as this was the most appropriate method for the present data set. Analyses were made according to the formula given in "The Estimation of Sensitivity and Specificity of Clustered Binary Data". Receiver operating characteristic (ROC) analysis of standardized uptake values (SUV) and needle biopsies compared to Gleason scores were performed by using SAS PROC LOGISTIC's ROC routine (SAS Institute Inc., Cary, North Carolina, USA). The SUV cut-off has been calculated by maximizing the Youden Index (maximizing the sum of sensitivity and specificity) excluding
values where sensitivity or specificity equals 100%. Maximizing the Youden Index method resulted in a SUVmean t1 cut-off of 3.8, however, excluding cases with sensitivity or specificity of 100%, the second highest Youden Index was chosen, resulting in a cut-off of 3.4. In order to quantify the overall staging concordance of FEC-PET/MRI vs. histology, a post-hoc kappa analysis was performed. The interpretation of the calculated kappa (.3871) is ‘fair’ according to D. Altman: Practical Statistics for Medical Research, 1991.


