Received: 26 August 2019 | Accepted: 10 November 2019

DOI: 10.1002/pros.23930

ORIGINAL ARTICLE

The Prostate WILEY

Ultrasound and magnetic resonance imaging for group stratification and treatment monitoring in the transgenic adenocarcinoma of the mouse prostate model

Stein-Martin T. Fagerland MD[1,2] | Deborah K. Hill PhD[2] | Annemieke van Wamel PhD[1] | Catharina de Lange Davies PhD[1] | Jana Kim PhD[2]

[1]Department of Physics, Norwegian University of Science and Technology, Trondheim, Norway

[2]Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, Trondheim, Norway

Correspondence
Catharina de Lange Davies, Department of Physics, NTNU, Høgskoleringen 5, 7491 Trondheim, Norway.
Email: catharina.davies@ntnu.no

Funding information
The Research Council of Norway, Grant/Award Number: 240316; Norwegian Cancer Society, Grant/Award Number: 6824920-2015; Liaison committee between the Central Norway Regional Health Authority and NTNU, Grant/Award Number: 90059700

Abstract

Background: The transgenic adenocarcinoma of the mouse prostate (TRAMP) is a widely used genetically engineered spontaneous prostate cancer model. However, both the degree of malignancy and time of cancer onset vary. While most mice display slowly progressing cancer, a subgroup develops fast-growing poorly differentiated (PD) tumors, making the model challenging to use. We investigated the feasibility of using ultrasound (US) imaging to screen for PD tumors and compared the performances of US and magnetic resonance imaging (MRI) in providing reliable measurements of disease burden.

Methods: TRAMP mice (n = 74) were screened for PD tumors with US imaging and findings verified with MRI, or in two cases with gross pathology. PD tumor volume was estimated with US and MR imaging and the methods compared (n = 11). For non-PD mice, prostate volume was used as a marker for disease burden and estimated with US imaging, MRI, and histology (n = 11). The agreement between the measurements obtained by the various methods and the intraobserver variability (IOV) was assessed using Bland-Altman analysis.

Results: US screening showed 81% sensitivity, 91% specificity, 72% positive predictive value, and 91% negative predictive value. The smallest tumor detected by US screening was 14 mm^3 and had a maximum diameter of 2.6 mm. MRI had the lowest IOV for both PD tumor and prostate volume estimation. US IOV was almost as low as MRI for PD tumor volumes but was considerably higher for prostate volumes.

Conclusions: US imaging was found to be a good screening method for detecting PD tumors and estimating tumor volume in the TRAMP model. MRI had better repeatability than US, especially when estimating prostate volumes.

KEYWORDS
cancer monitoring, preclinical, prostate cancer, prostate volume, repeatability, tumor volume

1 | INTRODUCTION

Prostate cancer (PCa) is the most common cancer in men in the European Union with estimated 450 000 new cases in 2018.[1] Even though overall survival is high, 107 300 men were estimated to die from PCa in 2018.[1] New treatments are required to improve survival, and good disease models are needed to select the most promising treatments in preclinical trials. Mouse cancer models are often used even though the mouse prostate differs from the human prostate in many respects. Unlike the human prostate, the mouse prostate does not have a fibrous capsule, is divided into multiple lobes, and is located in the abdominal cavity. The most commonly used mouse models in preclinical PCa studies are immunodeficient mice with subcutaneous xenografts grown from human cancer cell lines.[2] Several factors make these models less representative of human cancer, for example, immunodeficiency, homogenous tumor genetics, and, in the case of subcutaneous models, ectopic placement of the tumor.[2] It is believed that genetically engineered models (GEMs), with a functional immune system and orthotopic spontaneous cancer onset, are superior models that more accurately mimic human disease. Despite its promise, the use of GEM is still limited because of large variations in the degree of malignancy and age at cancer onset, which lead to high variation within treatment groups and planning experiments can be challenging.

The transgenic adenocarcinoma of the mouse prostate (TRAMP) is a widely used spontaneous PCa model.[3,4] TRAMP mice develop cancer gradually from normal cells that transition through prostatic intraepithelial neoplasia to cancer.[5-7] Because of neoplasia, the TRAMP prostate outgrows wild-type mouse prostates by 12 weeks of age. The TRAMP prostates continue growing as mice age, while wild-type mouse prostate volumes stabilize by 24 weeks of age.[8] However, the disease progression is heterogenous, and a subgroup develops fast-growing poorly differentiated (PD) tumors from around 12 weeks of age.[9] After PD tumor detection, mice often need to be euthanized within a few weeks because of tumor burden affecting animal welfare, whereas mice without PD tumors can live much longer, often up to 1 year of age.[10] Consequently, the TRAMP model displays distinct phenotypes with different disease progression, which can make it challenging to determine treatment timepoints and to measure treatment response.

Variation in the cancer phenotype and age at cancer onset in the TRAMP model requires strategies to reduce the variation in outcome measurements. One strategy is starting the treatment when mice are young and the variation between individual mice is low. In such studies, treatment starts before cancer onset and the studies are often called chemopreventive. Another strategy is to use longitudinal outcome measurements acquired with in vivo imaging, enabling stratification of the mice in different treatment groups based on the phenotype and the use of longitudinal statistical methods.

A number of outcome measurements are used to compare the efficacy of treatments in the TRAMP model. Methods based on histology are most commonly used and several grading systems based on hematoxylin and eosin-stained sections have been developed for the TRAMP model[6,7,10-13] and are often combined with immunohistochemistry markers for proliferation, apoptosis, and others. Often, tumor or prostate size is used to measure the treatment effect. The measurement is usually done ex vivo, since the prostate and prostate tumors are inaccessible for caliper measurements. Less frequently, imaging techniques are used to assess disease progression or treatment response in vivo. For PD mice, tumor volume is sometimes used to measure tumor burden.[9] However, most TRAMP mice do not develop PD tumors and an alternative treatment efficacy outcome such as prostate volume must be used.[14,15] It has previously been shown that the prostate volume of the TRAMP background strain C57BL/6 does not usually grow past 40 mm^3, while TRAMP mice can reach a volume of around 100 mm^3 when they are 25 weeks of age.[8] The change in prostate volume can, therefore, be attributed to neoplasia of the prostate in the TRAMP and be a marker of disease progression. However, the growth is not always caused by malignant disease and might not indicate aggressiveness or metastatic potential.

We hypothesize that in vivo imaging can be used to overcome challenges posed by the TRAMP model, by offering longitudinal outcome measurements and treatment group stratification based on phenotype. This could enable the use of more powerful longitudinal statistical analysis and reduce the number of mice needed per treatment group. Examples of longitudinal measurements are volume change over time, volume doubling time, and tumor-free survival. In this study, we investigated the performance of ultrasound (US) and magnetic resonance imaging (MRI) for imaging pathology of the prostate. Both imaging modalities are widely available in preclinical imaging facilities. Compared with MRI scanners, US scanners are generally less expensive to purchase, operate, and require less infrastructure. Moreover, US image acquisition is often quicker than MRI. Consequently, we were particularly interested in examining whether US imaging could be used to detect and measure changes in the prostate in the TRAMP model.

Thus, this study addresses two of the main challenges working with the TRAMP model; the unpredictable onset of PD tumors, and longitudinal monitoring of cancer progression or treatment response. First, we investigated the PD tumor screening performance of US imaging using MRI or gross pathology (in two cases) for verification. Second, we estimated PD tumor volume and compared agreement and intraobserver variability (IOV) for US imaging and MRI. Finally, we performed prostate volume measurements for non-PD mice and compared the agreement and IOV for US imaging, MRI, and histology.

2 | MATERIALS AND METHODS

2.1 | Mice

Mice were bred from an in-house colony established at the Norwegian University of Science and Technology (NTNU, Trondheim, Norway), with TRAMP mice originally purchased from the Jackson Labs in 2012. Initially, TRAMP mice homozygous for the TRAMP mutation were used in our experiments, but after a literature review, we changed to breeding heterozygous TRAMP mice in accordance with the majority of TRAMP studies. The in-house colony consisted of TRAMP mice homozygous for the TRAMP mutation. Heterogenous mice were bred using homozygous TRAMP females and C57BL/6 males acquired from

the Jackson Labs. The presence of TRAMP mutation in the offspring was verified by a polymerase chain reaction. PD tumor screening and volume estimation were performed on mice heterozygous for the TRAMP mutation and prostate volume were calculated for TRAMP mice homozygous for the TRAMP mutation. All mice were housed under specific pathogen-free conditions, in groups of one to six in ventilated cages (Model 1284L; Tecniplast) at temperatures from 21°C to 23°C, with 45% to 60% relative humidity, 70 air changes per hour and ad libitum access to food and sterile water.

2.2 | Ethics

All animal experiments were approved by the Norwegian Animal Research Authority and Norwegian Food Safety Authority.

2.3 | US imaging protocol and volume estimation

PD tumor screening and measurement of tumor volume were performed using a FUJIFILM Visualsonics Vevo 3100 scanner with an MX550D probe with 40 MHz center frequency, $30 \times 30\ \mu m^2$ in-plane resolution and 80 µm through-plane resolution. Images were acquired every 76 µm using a 3D-motor. Imaging for prostate volume estimation was done with a FUJIFILM Visualsonics Vevo 2100 scanner with an MS550D probe with 40 MHz center frequency, $40 \times 40\ \mu m^2$ axial resolution, and 90 µm lateral resolution. Images were acquired every 32 µm using a 3D-motor. The reason for the system change was an upgrade. Mice were kept under anesthesia during imaging using 1.5% to 2.5% isoflurane with 0.5 L/min 5:1 air:O_2 mix. Mice were restrained in the supine position by taping the legs to the imaging stage. Respiration was monitored visually, and anesthesia was adjusted to a respiration rate around 80 breaths per minute. Before imaging, the abdomen was shaved with an electric razor and depilation cream was used to remove remaining fur. US gel was applied before imaging. The US transducer was placed over the lower abdomen with an axial orientation. The bladder was easy to locate because of its size and anechoic urine content. After identifying the bladder, the transducer was swept caudally until the urethra was identified. The prostate was identified as a heterogenous, hypoechoic structure ventral, lateral and often dorsal to the urethra. Gain and dynamic range were adjusted to see the prostate clearly while maintaining contrast to surrounding tissue. During PD tumor screening, the transducer was pressed manually against the pelvic region to detect tissues that were stiffer than the prostate. PD tumors were identified as spherical, hypoechogenic, with a heterogenous contrast and were less compressible than surrounding tissue. After the extent of the prostate had been identified, a 3D-scan was acquired with the 3D-motor. Image stacks were imported into the Vevo LAB (FUJIFILM Visualsonics) software, processed with the "load into 3D function," exported as "TIFF for 3D Volume Slice (*.tif)" and opened in FIJI ImageJ. A region of interest (ROI) was drawn around the PD tumor or prostate depending on the phenotype. For mice without PD tumor, ROIs were drawn over the ventral, dorsal, and lateral prostate lobes for approximately every 5th to 10th frame and interpolation was performed between frames. In the cranial direction, segmentation stopped when the seminal vesicle appeared, visualized as hypo or anechoic cystic septate structures. The volume of the ROIs in the image stack was calculated based on voxel size.

2.4 | MRI protocol and volume estimation

MRI was performed on a Bruker 7T Biospec 70/20 Avance III. Axial T2-weighted images were acquired using a RARE spin-echo sequence with fat suppression. For PD tumor detection and size measurement, several different coil setups and sequence parameters were used due to optimization at the beginning of the study: (a) four PD tumors were verified using an 86 mm volume resonator coil for RF transmission and reception with TE 58.5 ms, TR 4000 ms, RARE factor 8, averages 6, in-plane resolution $0.2 \times 0.2\ mm^2$, slice thickness 0.8 mm and acquisition time 3.7 minutes; (b) three PD tumors were verified using an 86 mm volume resonator coil for RF transmission and a phased array rat brain surface coil for RF reception with TE 58.5 ms, TR 4000 ms, RARE factor 8, averages 12, in-plane resolution $0.15 \times 0.15\ mm^2$, slice thickness 0.6 mm and acquisition time 5.2 minutes; (c) nine PD tumors were verified using an 86 mm volume resonator coil for RF transmission and a phased array mouse heart surface coil for RF reception with TE 58.5 ms, TR 5000 ms, RARE factor 6, averages 6, in-plane resolution $0.1 \times 0.1\ mm^2$, slice thickness 0.4 mm, and acquisition time 10 minutes. The third setup was preferred since the mouse heart coil gave the best signal from the prostate area compared with the volume resonator coil and the rat brain coil. All MRI PD tumor volume measurements were done with the third setup. For prostate volume measurements, axial T2w RARE images were acquired using an 86 mm volume resonator coil for RF transmission and a phased array mouse heart surface coil for RF reception with TE 36 ms, TR 5500 ms, RARE factor 8, averages 5, in-plane resolution $0.1 \times 0.1\ mm^2$, slice thickness 0.33 mm, and acquisition time 7 minutes. The TE was changed from 36 to 58.5 ms to increase the contrast between prostate tissue and seminal vesicle.

Mice were kept under anesthesia with 1.5% to 2.5% isoflurane with 0.5 L/min 5:1 air:O_2 mix. Respiration was monitored with a pressure sensor connected to a physiological monitoring system (SA Instruments) and anesthesia adjusted to a respiration rate around 80 breaths per minute. Mice were restrained in the scanner bed in a prone position using tape across the lower back. The imaging volume was selected by using a quick localizer scan and a quick low-resolution scan in both axial and coronal orientation to identify the extent of the tumor or prostate. PD tumors were identified as spherical with a homogenous contrast and hyperintense signal compared with the surrounding tissue. Only the bladder displayed a higher signal than PD tumors. Prostate tissue was identified based on its heterogenous signal and location mainly ventral, lateral, and dorsal to the urethra. MRI images were exported as DICOM files from Paravision 6.0.1 and loaded into FIJI ImageJ where ROIs were drawn manually on all volume slices and ROI volumes estimated.

2.5 | Histology protocol and volume estimation

Eleven TRAMP mice were imaged by US and MRI the same day and were euthanized immediately after imaging to harvest the prostates.

The prostate was harvested together with the seminal vesicle, bladder, and urethra and fixed in 4% formaldehyde and paraffin-embedded before sectioning with a microtome (Leica RM2255). Before embedding, the prostate was oriented using the bladder, ductus deference, and urethra as landmarks, to align histology sections to axial images from US and MRI to be able to compare the three diagnostic modalities. From all eleven prostates, five sections (5 µm thick) were collected every 200 µm and stained with hematoxylin-erythrosine-saffron (HES). Saffron stains collagen fibers, and HES staining is standard at our university's core facility. However, adding saffron did not help distinguishing prostate from surrounding tissue. Every other slide was scanned using brightfield with a Zeiss LSM810 microscope using a ×2.5 objective with a numerical aperture of 0.085. Multiple images were acquired to cover each section and images were stitched together using the Zeiss Zen software. Files were opened in Zeiss Blue software and one ROI was drawn around the ventral, dorsal, and lateral prostate lobe in each section. ROIs were drawn based on characteristic histological features.[16] After ROIs were drawn, the area of the ROIs was multiplied with the distance between slides and summed to estimate the prostate volume.

2.6 | Tumor screening

US screening for PD tumors was performed in 81 mice every third week in the age range from 16 to 24 weeks (Figure S1). The total number of US screens per mouse ranged from one to three. After PD tumor verification with MRI, no further US imaging was performed. MRI was used to verify the US findings in 72 mice and gross pathology in two mice. Seven mice were excluded since they died or had to be euthanized before MRI at week 25 due to kidney tumors or reduced activity as a sign of illness. Initially, US screening started at week 20, but after two mice needed to be euthanized because of large tumors, screening was moved to week 17 ± 1 week. These two mice were euthanized immediately due to animal welfare requirements, and US findings verified by gross pathology.

Mice with PD tumors detected by US screening were scanned with MRI the following day to verify the result. Mice with a negative US screen or a MRI verified false-positive result, were imaged again using US after 3 to 4 weeks. At 23 to 26 weeks of age, all mice were imaged with MRI to verify earlier US findings.

US screening performance was assessed by estimating sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Sensitivity was defined as a number of true-positive/(true-positive + false-negative), specificity as true-negative/(true-negative + false-positive), PPV as true-positive/(true-positive + false-positive) and NPV as true-negative/(true-negative + false-negative). PD tumor growth rate was assessed in four mice that were imaged weekly from week 19 or 20 of age until tumor burden necessitated euthanasia. Tumor doubling time was estimated by fitting tumor growth data to an exponential function $V(t) = V_0\exp(bt)$. The curve fitting toolbox in MATLAB R2018b was used to estimate the rate constant b and the doubling time was given by $t_d = \ln(2)/b$.

2.7 | Imaging modality comparison and IOV of volume estimates

PD tumor and prostate volume estimates were based on the sum of voxel volumes within ROIs containing PD tumors or prostate tissue. For mice without PD tumor, ROIs were limited to the ventral, dorsal, and lateral lobes of the prostate, since the anterior prostate lobe, can be difficult to distinguish from the seminal vesicle on images. ROIs were drawn in the same image stack by the same observer two times at least 1 week apart and the repeated measurements were used to assess IOV and to compare imaging modalities using the methods described by Bland and Altman.[17] Volume estimations were done by SM Fagerland (MD), who had limited experience with US imaging, MRI, and histology at the start of the study. US training was supplied in two sessions by application specialists from FUJIFILM VisualSonics, MRI training was given by researchers with extensive experience with the TRAMP model and identification of the prostate on histological sections was based on the papers by Berman-Booty[7] and Oliveira.[16] Agreement between imaging modalities was visualized using a scatter plot and a Bland-Altman plot, where the difference between the modalities is plotted on the vertical axis and the mean of the results obtained by the two modalities on the horizontal axis. Because the magnitude of the difference is dependent on the size measured, we used percentage difference from the mean on the vertical axis to correct for the size-dependent variation.[18] The mean of the differences between the repeated measurements was calculated and represented the bias or the systematic difference between the measurements. The standard deviation (SD) of the differences between repeated measurements was also calculated. The agreement between modalities and IOV was quantified using limits of agreement as defined by Bland-Altman. The limits of agreement are calculated as bias ±1.96 SD. When the modalities were compared, the mean of two repeated measurements was used and the calculation of limits of agreement was corrected according to Bland-Altman.[17] The precision of the estimated bias and limits of agreement were calculated using 95% confidence intervals (CIs), as described in the paper by Giavarina.[18] Eleven mice with PD tumors were used to compare the tumor volume obtained by US imaging and MRI and to investigate the IOV of US and MRI volume estimations. Similarly, 11 mice with age span 21 to 31 weeks without PD tumors were used to compare prostate volume measurements obtained by US imaging, MRI, and HES sections and to investigate the IOV for each modality.

3 | RESULTS

3.1 | Characterization of TRAMP phenotypes

The TRAMP mice exhibited two different phenotypes. All mice displayed a hyperplastic prostate with atypia, but 22% developed a fast-growing PD prostate tumor in addition. No TRAMP mouse developed more than one PD tumor. Representative HES sections from the two phenotypes are shown in Figure 1. The PD tumors have few, sparse glands, and display anaplastic sheets of pleomorphic cells. The prostate tissue is generally

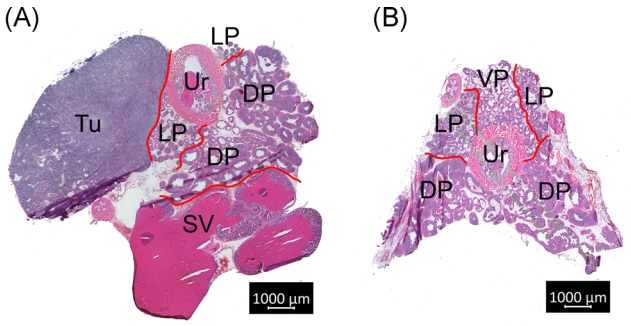

FIGURE 1 A, HES section from a TRAMP mouse prostate with a PD tumor of approximately 200 mm^3 measured by MRI. B, HES section of prostate with a high degree of prostatic intraepithelial neoplasia in all lobes and especially the dorsal prostate. DP, dorsal prostate; HES, hematoxylin-erythrosine-saffron; LP, lateral prostate; MRI, magnetic resonance imaging; SV, seminal vesicle; TRAMP, transgenic adenocarcinoma of the mouse prostate; Tu, tumor; Ur, urethra; VP, ventral prostate [Color figure can be viewed at wileyonlinelibrary.com]

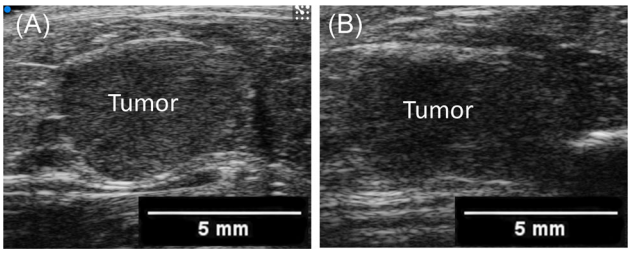

FIGURE 3 A, An US image of a PD tumor from a 20-week-old mouse acquired under compression and shows a clear border between the tumor and surrounding tissue. B, The same tumor with low compression during the same image session where the border is less well-defined. The tumor had a maximum diameter of 4.6 mm and volume 39.4 mm^3 estimated by US. Based on MRI the volume was estimated to be 56.7 mm^3. MRI, magnetic resonance imaging; PD, poorly differentiated; US, ultrasound [Color figure can be viewed at wileyonlinelibrary.com]

characterized by cells filling the lumen of glands, herniation of cells into the smooth muscle layer, and hyperplastic epithelium.

3.2 | PD tumor screening

The PD prostate tumors appeared relatively spherical and homogenous compared with surrounding tissue on both US and MR images and were often located in the ventral or lateral lobe of the prostate (Figure 2). Left-right image asymmetry could also be a sign of PD tumor, since the normal prostates display a left-right symmetry in axial images. US imaging and MRI differed in PD tumor signal intensity compared with surrounding tissue; PD tumor signal intensity was relatively low for US imaging and relatively high for T2-weighted MR images. Tumors smaller than 3 mm often had a heterogenous contrast and unclear border on US images. Compression could be used to enhance the contrast between tissues, as shown in Figure 3. The border between the tumor and surrounding tissue appeared more distinct when the tissue was compressed, as PD tumors were not as compressible as surrounding prostate tissue.

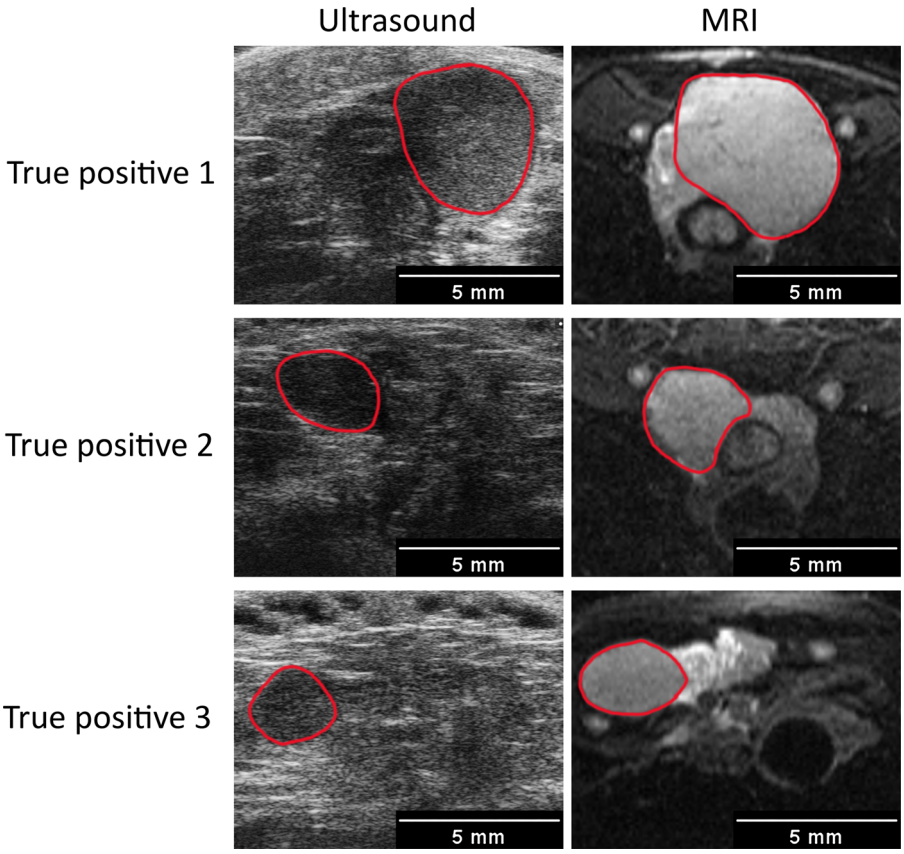

FIGURE 2 PD prostate tumors (outlined in red) imaged with US in the left column and MRI in the right column. The three mice were 17 weeks old when imaged. The tumors appear relatively spherical and homogenous compared with surrounding tissues. The signal intensity of PD prostate tumors is relatively low on US imaging and relatively high on T2-weighted MRI, compared with the surrounding tissue. MRI, magnetic resonance imaging; PD, poorly differentiated; US, ultrasound [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 US screening performance in TRAMP mice verified by MRI (n = 72) and gross histology (n = 2)

Sensitivity	81%	True-positive	n = 13
Specificity	91%	False-positive	n = 5
Positive predictive value	72%	False-negative	n = 3
Negative predictive value	91%	True-negative	n = 53
Prevalence of prostate PD tumors	22%		

Abbreviations: MRI, magnetic resonance imaging; PD, poorly differentiated; TRAMP, transgenic adenocarcinoma of the mouse prostate; US, ultrasound

US imaging performed well in PD tumor screening. US imaging was performed one to three times per mouse in the age range 16 to 24 weeks. This resulted in 81% sensitivity, 91% specificity, 72% PPV, and 91% NPV (Table 1). Sixty-six cases were classified correctly as true-positive (13 mice) or as true-negative (53 mice). Eight cases were misclassified as false positives (5 mice) or false-negative (3 mice). The false negatives were detected with MRI 3 weeks after the last US screening and volumes were in the range of 21 to 93 mm^3 measured with MRI. In four out of five false-positive cases the ventrolateral area of the prostate was misinterpreted as tumor (Figure 4). Misinterpretation occurred since the prostate sometimes shares characteristics with PD tumors by appearing spherical during compression or a tilted mouse position can make the prostate appear asymmetrical. Twenty-two percent of the mice (16 of 74) developed PD tumors by 25 weeks of age.

Figure 5 shows mouse age and tumor volume at the time of PD prostate tumor detection either by US imaging or MRI. Volumes were estimated using MRI or US. US screening detected PD tumors as small as 14 mm^3 measured with MRI and a maximum diameter of 2.6 mm when measured with US. PD tumor volume varied widely in mice of the same age, ranging from 14 to 420 mm^3 in week 17 and 16 to 3100 mm^3 in week 20. No PD tumors were detected at week 16 or 18. Four of the mice with PD tumors were imaged weekly by MRI until they reached a tumor diameter of more than 1.5 cm. The growth data for these four mice are plotted in Figure 5. The growth data were fitted to an exponential function and all R^2 values were greater than 0.99. On the basis of the exponential function, PD tumor doubling time was 4.0, 4.2, 5.3, and 5.9 days.

3.3 | Comparing tumor volume estimated by US imaging and MRI

PD tumor volumes (Table S1) estimated from US images correlated well with, but were generally smaller than MRI measurements, as

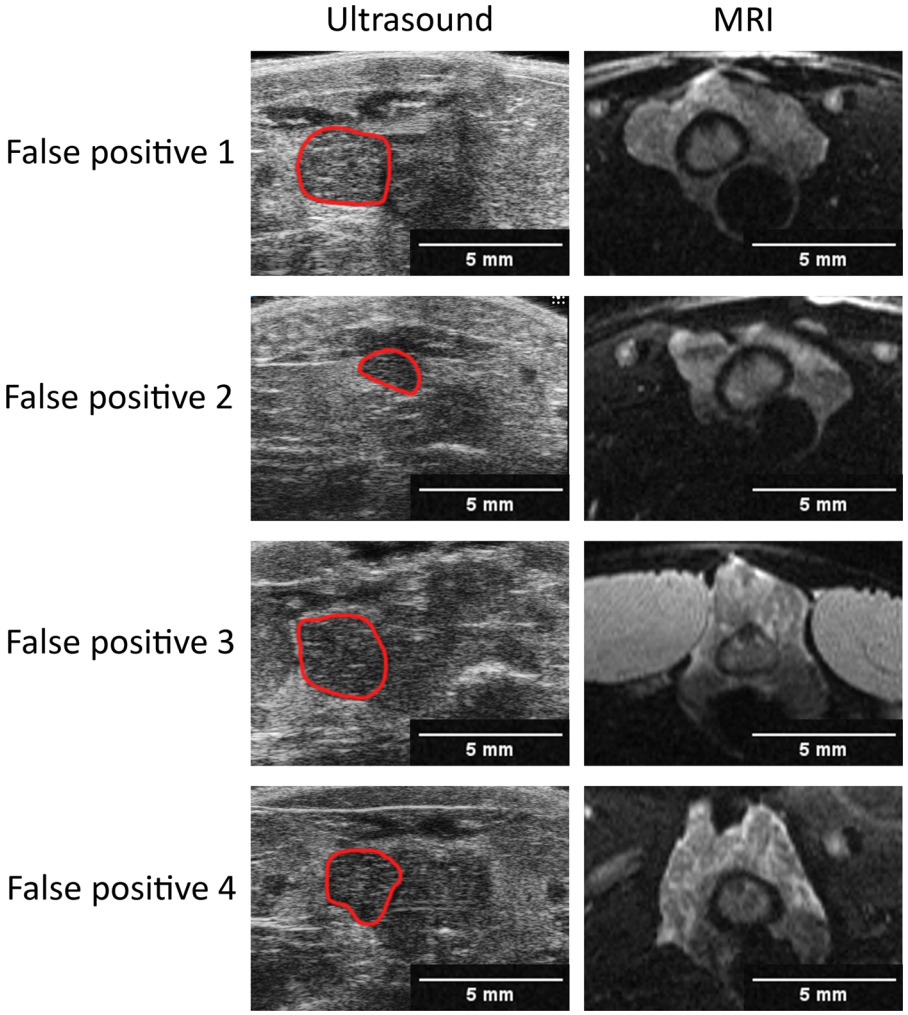

FIGURE 4 Images of the TRAMP model where US screening gave false-positive results. Corresponding US and MR images are shown in the left and right columns, respectively. Structures misinterpreted as PD tumors on US are outlined in red. Misinterpretation of the ventrolateral aspect of the prostate on US images led to four out of five false positives. MR, magnetic resonance; PD, poorly differentiated; TRAMP, transgenic adenocarcinoma of the mouse prostate; US, ultrasound [Color figure can be viewed at wileyonlinelibrary.com]

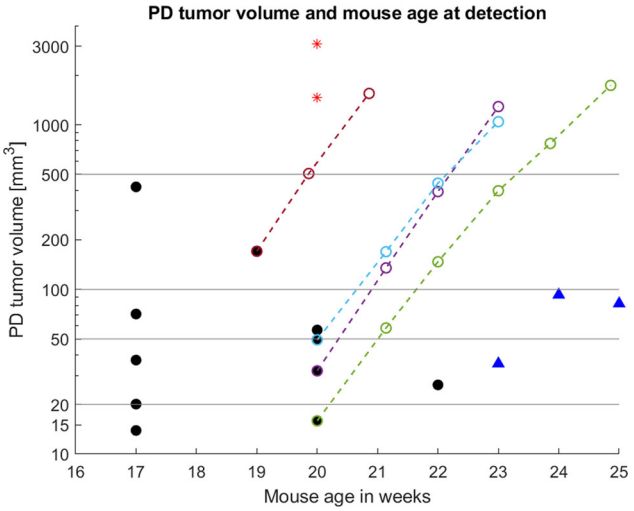

FIGURE 5 PD tumor volumes (n = 16) are plotted on a logarithmic scale on the vertical axis and mouse age at PD tumor detection on the horizontal axis. Black circles (n = 11) represent individual tumors detected with US and volumes were estimated by MRI. Red asterisks (n = 2) represent individual tumors for which volumes were estimated by US. Blue triangles (n = 3) indicate individual tumors detected and measured by MRI but not US (false negatives). The colored, dashed lines show growth data for four PD tumors. All growth data fitted an exponential function with $R^2 > 0.99$. MRI, magnetic resonance imaging; PD, poorly differentiated; US, ultrasound [Color figure can be viewed at wileyonlinelibrary.com]

demonstrated from the log-log scatter plot in Figure 6A. The tumor volumes for the 11 tumors varied from 8.1 to 587 mm^3 when estimated with US imaging and 8.5 to 721 mm^3 when estimated with MRI. To obtain more detailed information about the difference in the volumes obtained by the two imaging modalities, the data were analyzed in a Bland-Altman plot (Figure 6B). The blue line indicates the systematic difference between the methods and shows that the tumor volumes measured by US imaging were on average 22% smaller than volumes measured by MRI, with a 95% CI of 5.4% to 38.9%. The red lines indicate the limits of agreement, which were −71.9% to 27.6% (Figure 6B).

3.4 | IOV in tumor volume estimation

IOV was slightly larger for US imaging than for MRI as shown on the Bland-Atman plot (Figures 7 and Tables S2, S3). The bias for US imaging and MRI was 2.6% (95% CI: −1.8%-7.1%) and −0.1% (95% CI: −3.8%-3.7%), respectively, indicating no systematic difference between the repeated measurements. The limits of agreement for volume measurements were −11.2% to 16.5% for US imaging and −11.1% to 10.9% for MRI.

3.5 | Comparing prostate volume measured by US imaging, MRI, and histology

Figure 8 shows prostate tissue ROIs for US imaging, MRI, and histology in three different mice (Tables S4-S6). The contrast between the prostate and surrounding tissues can be low, especially for US imaging. The border between the prostate and seminal vesicle is the most difficult to distinguish in US imaging and MRI, because of the uneven nature of the border and relatively low contrast. On HES images, the dorsal and anterior lobe could sometimes be mistaken for each other. Normally the lobes can be distinguished based on the epithelial lining of the gland, which is more columnar for the anterior gland, and the secretion, which is more eosinophilic in the anterior

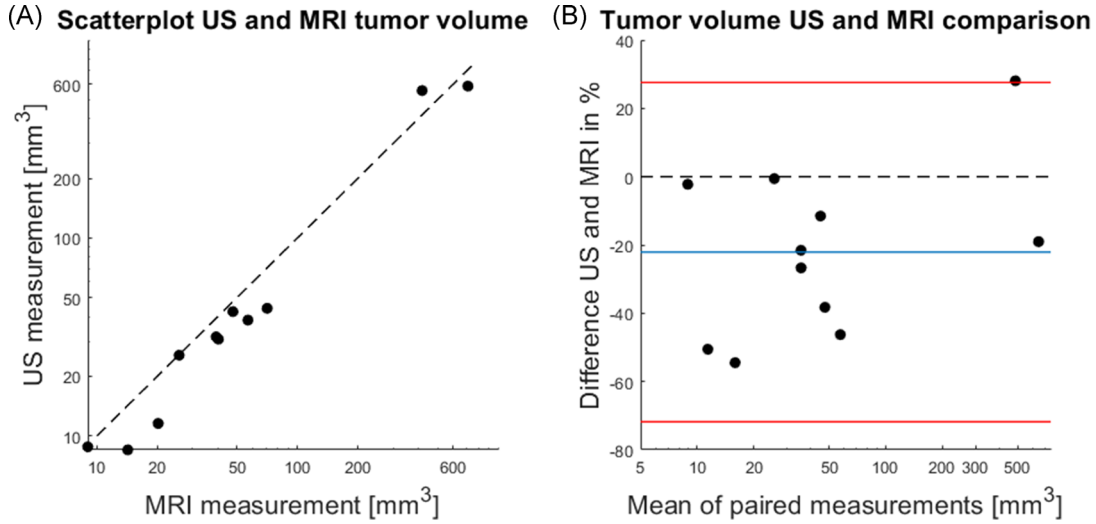

FIGURE 6 A, Log-log scatter plot comparing tumor volumes obtained by US imaging and MRI. The dashed line shows the line of equality. Circles represent the mean of paired PD volumes estimated by MRI vs PD volume estimated by US imaging. B, Bland-Altman plot comparing tumor volume estimated by US imaging and MRI. The vertical axis shows the percentage difference between volumes obtained by US imaging and MRI, and the horizontal axis shows the mean of the volumes from US imaging and MRI on a logarithmic scale. The blue line shows the mean of the differences, indicating that the tumor volumes measured by US imaging are on average 22% lower than MRI measured volumes. The red lines show the limits of agreement. MRI, magnetic resonance imaging; PD, poorly differentiated; US, ultrasound [Color figure can be viewed at wileyonlinelibrary.com]

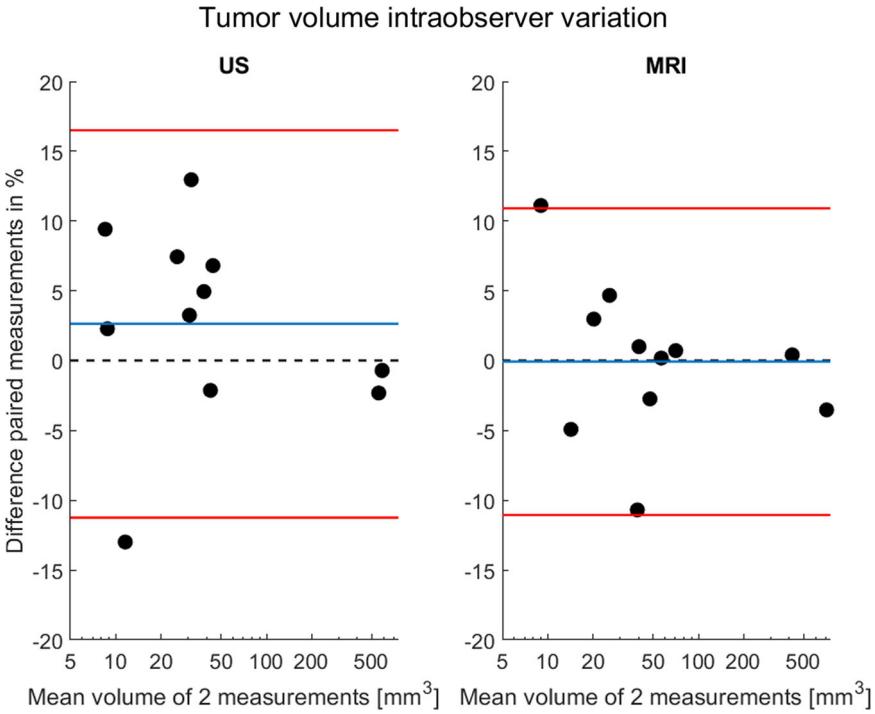

FIGURE 7 Bland-Altman plots of IOV in tumor volume obtained from US imaging and MRI. The vertical axis shows the percentage difference between two repeated measurements, and the horizontal axis shows the mean of the two repeated measurements (logarithmic scale) measured by US (left) or MRI (right). The blue lines indicate the mean of the differences and the red lines show the limits of agreement. Volume estimates from MRI have lower IOV than US imaging. IOV, intraobserver variability; MRI, magnetic resonance imaging; US, ultrasound [Color figure can be viewed at wileyonlinelibrary.com]

gland. However, sometimes neoplastic epithelium fills the glands and the hallmarks identifying the glands can disappear.

Prostate volume estimations from US images were, in general, much smaller than estimations from histology and MRI. MRI volume estimations were closer to histology volumes with a small overestimation, as demonstrated from the log-log scatter plot in Figure 9A. The prostate volumes from 11 mice varied from 19.6 to 55.8 mm^3 on US imaging, 33.6 to 97.4 mm^3 on MRI, and 31.8 to 70.9 mm^3 on histology. Comparisons between US and histology, US and MRI, and MRI and histology are plotted in a Bland-Altman plot to visualize and quantify the agreement

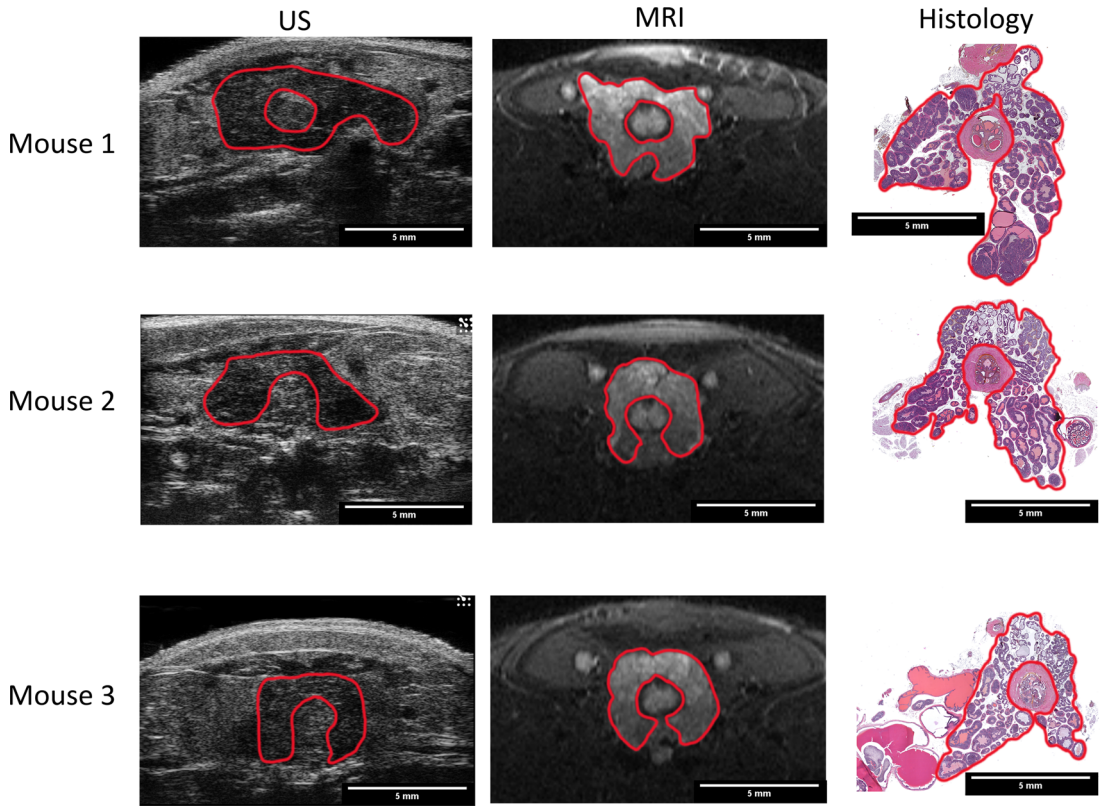

FIGURE 8 Segmentation of prostate for volume estimation in three representative mice with US imaging, MRI and histology. Prostate tissue is outlined in red. MRI, magnetic resonance imaging; US, ultrasound [Color figure can be viewed at wileyonlinelibrary.com]

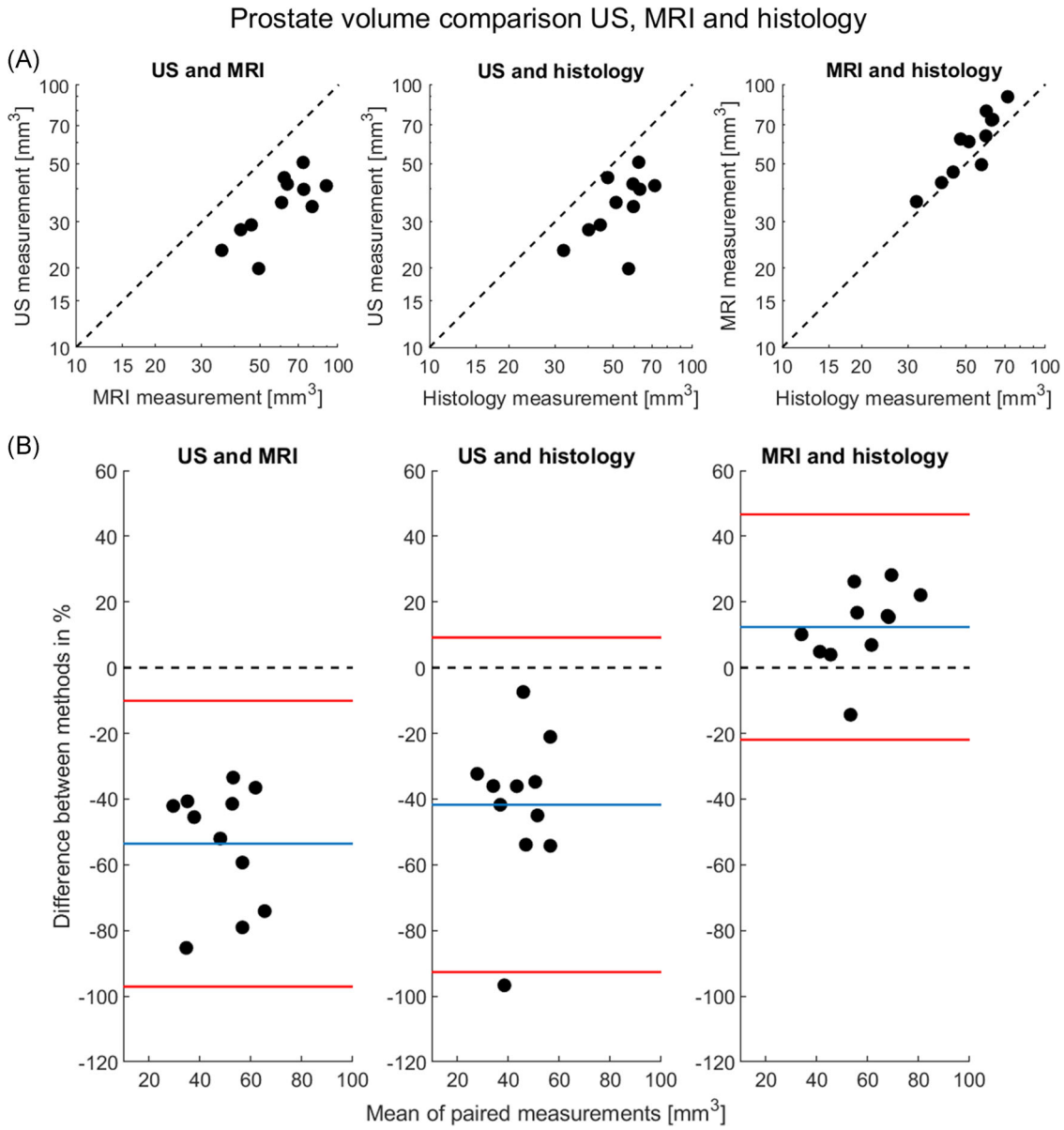

FIGURE 9 A, Log-log scatterplots comparing prostate volumes obtained from US imaging, MRI, and HES section scans. The dashed line shows the line of equality. Each circle represents the prostate volume of a different mouse. B, Bland-Altman plots comparing volume estimations from US, MRI, and histology. Percentage differences between measurements are plotted on the vertical axis and the mean of the measurements is plotted on the horizontal axis. The blue lines show the mean of the differences, and the red lines show the limits of agreement. HES, hematoxylin-erythrosine-saffron; MRI, magnetic resonance imaging, US, ultrasound [Color figure can be viewed at wileyonlinelibrary.com]

(Figure 9B). The limits of agreement between US and histology were −92.7% to 9.2% with a bias of −41.7%, the limits of agreement between US and MRI were −97.1% to −10.1% with a bias of 53.6% and the limits of agreement between MRI and histology were −21.9% to 46.6% with a bias of 12.4%.

3.6 | IOV in prostate volume

The IOV for prostate volume measurement was much larger for US imaging compared with MRI and histology, whereas the IOV for prostate volumes obtained by MRI and histology were almost the same as shown on the Bland-Altman plot (Figure 10 and Tables S7-S9). The mean of the differences for US imaging, MRI, and histology was not significantly different from 0, indicating no systematic difference between the repeated measurements. The limits of agreement were −35.1% to 53.9% for US imaging, −24.3% to 20.2% for MRI and −26.8% to 19.0% for histology.

4 | DISCUSSION

Since the establishment of the TRAMP model in 1995,[19] the model has been widely used to study PCa progression and to develop new therapies and strategies to prevent or treat PCa. However, working with the

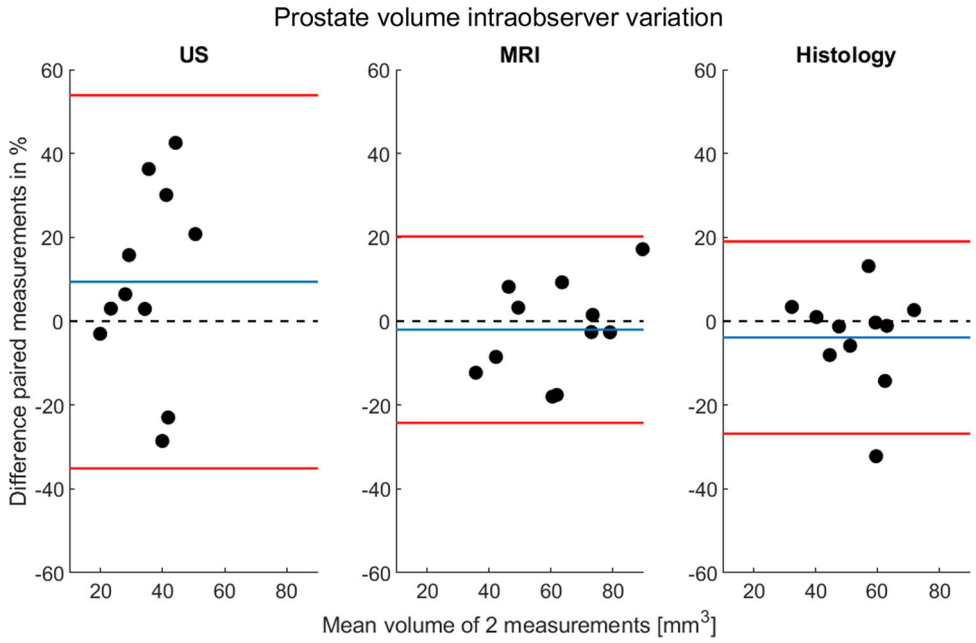

FIGURE 10 Bland-Altman plots showing IOV for volume estimation based on US imaging, MRI, and histology. The vertical axis is percentage differences between two repeated measurements and the horizontal axis is the mean of the two repeated measurements. The blue lines show the mean of the differences. The red lines show the limits of agreement. MRI and histology show comparable IOV, while US has a larger IOV. IOV, intraobserver variability; MRI, magnetic resonance imaging; US, ultrasound [Color figure can be viewed at wileyonlinelibrary.com]

TRAMP model is challenging because of unpredictable disease onset characteristics and variation in cancer malignancy. While histology has been the standard method to evaluate treatment response, imaging can improve treatment evaluation by enabling longitudinal measurements or stratifying TRAMP phenotypes before commencing treatment. Stratification reduces the number of mice needed in preclinical studies through a reduction in the variation in treatment response between groups. Most TRAMP mice do not develop solid tumors during their life span, but show the cancerous transformation of the prostate and increase in prostate volume; a smaller subgroup develops PD tumors, which causes a large variation in treatment outcomes. We demonstrated that US imaging, being a low-cost and broadly available imaging modality, could be used to screen TRAMP mice for PD tumors. In addition, PD tumor volumes obtained by US and MRI were compared, as well as prostate volumes obtained by US imaging, MRI, or histology.

4.1 | PD tumor screening

Stratification of mice into different treatment groups based on the disease phenotype is important to reduce variation in treatment response within treatment groups and to reduce the number of mice needed in preclinical trials. US imaging every third week, up to three times in total, was found to detect PD tumors in mice aged 16 to 24 weeks with 81% sensitivity and 91% specificity, demonstrating that US imaging can be a valuable tool for detecting PD tumors. For the false-negative cases, the volumes were too small for detection by US imaging and the validation with MRI performed 3 weeks later revealed volumes in the range of 21 to 93 mm^3. Thus, at the time of US imaging, the tumors were most likely below 2 mm in diameter. Misclassification of prostates as false positives occurred because the TRAMP prostate and PD tumors sometimes share characteristics, for example, compression can make the prostate appear spherical and a tilted image plane can make the prostate appear asymmetrical.

The screening performance could be improved by repeating US weekly. Wirtzfeld et al[20]achieved a sensitivity of 95% and specificity of 100% when screening 33 mice from the spontaneous PCa model PSP-TGMAP with US imaging weekly from 16 to 32 weeks of age. In that study, three investigators reached consensus over the presence or absence of tumors while having previous imaging data available. Together with the higher frequency of imaging sessions compared with our study, this led to an improved screening performance. Other factors that can contribute to the difference are the use of a different mouse model, different US set-up and fewer mice in the Wirtzfeld study. A limitation of both our and Wirtzfeld's study is that negative US screenings were not immediately validated by MRI or histology. In our study, mice with negative US screening were screened up to three times before validation and in Wirtzfeld's study, they were screened even more often.

The smallest tumor we detected at week 17 during US screening had a volume of 14 mm^3 measured on MRI and a maximum diameter of 2.6 mm on US, which was similar to the study by Wirtzfeld who reported tumor detection down to 2.4 mm. Degrassi et al[9] detected PD tumors with a diameter of 1 to 2 mm using MRI from week 12 of age, and the tumor doubling time for the PD tumors was 5 days, consistent with the doubling times in our study which were from 4.0 to 5.8 days. Several factors can influence the detection limit for PD tumors. In addition to tumor volume, the detection limit will depend on how close to the skin the tumor is located, if the tumor affects image symmetry and the tumor contrast to surrounding tissue.

When determining the mouse age at the start of screening and intervals between screening, important considerations are mouse age at PD tumor debut and PD tumor growth rate as well as the aim of the study to be performed. If a study requires small initial PD tumor volumes or one wants to follow tumor development closely, screening should start early (eg, at 16 weeks of age) and occur frequently (for example weekly). If the goal is to separate PD tumor mice from a population of TRAMP mice without tumors, screening can start later and be done less frequently. Knowledge of the tumor growth rate for a particular TRAMP strain might be used to extrapolate the growth curve to determine the start of screening and screening frequency to select mice with PD tumors of a certain volume.

4.2 | PD tumor measurement agreement and IOV

The repeatability of the volume measurements depends on the ROI placement around the tumor, which depends on how distinct the border between the tumor and surrounding tissue is. During US imaging, compression was used to improve the contrast between tissues and make borders more distinct, especially for tumors smaller than 3 mm in diameter. Large tumors often show a mass effect, compressing the tissue around it, giving a good contrast on US imaging without external compression. The agreement between US imaging and MRI was low, and the mean difference between measurements was 22% smaller with US imaging. Compression during US imaging could change the shape of the tumor, causing an apparent smaller volume especially for the smaller tumors where compression was done more often. Linxweiler et al[21] compared US and MRI estimated volumes for orthotopic prostate tumors and they also found that US imaging resulted in smaller volumes than obtained by MRI. Their US measurements were on average 4.85 mm^3 smaller than the MRI measurements and the limits of agreement were around −120 and 120 mm^3.[21] If we report our findings in absolute values rather than relative, they are very comparable to Linxweiler with a bias of 6.2 mm^3 and limits of agreement from −122 to 120 mm^3 (Figure S2). We quantified the difference with relative values instead of absolute values, since the variation depends on the tumor size and limits of agreement based on absolute values will be more dependent on larger tumors. In addition, one of the main benefits of using in vivo imaging is the longitudinal measurements quantify changes over time, which is a relative measurement, making relative measurements more relevant than absolute values.

The IOV for tumor volume estimation was slightly lower for MRI than US. The slightly better IOV for MRI could be because of a more distinct border and stronger contrast between tumor and the surrounding tissue compared with US imaging. This variation in measurement should be considered when estimating the number of animals needed for a study.

4.3 | Prostate volume measurement agreement and IOV

Prostate volumes were estimated by US imaging, MRI and histology. This is to our knowledge the first time a study comparing the three methods has been performed in TRAMP mice. We included the ventral, lateral, and dorsal prostate lobes in the volume estimate, since the TRAMP mutation affects all prostate lobes. The anterior lobe was not included in the volume estimate because it was hard to determine the extent of it by US imaging. Defining the ROI around the prostate in US and MR images can be difficult due to an unclear border between prostate lobes and surrounding tissue, especially between the dorsal prostate and the seminal vesicle. Volume estimates based on US imaging were 42% smaller than volumes estimated by histology and 54% smaller than volumes estimated from MRI. MRI and histology had a much better agreement. Compression during US imaging could have been a contributing factor for the low agreement between volumes obtained by US imaging and the other methods. During compression, the prostate changes shape and parts of it could have moved out of the imaged volume. The prostate volume measured by histology is expected to be less than volumes measured in vivo, because of shrinking during tissue fixation and subsequent tissue processing. The degree of shrinkage depends on both tissue and the processing steps. In kidney and prostate tumors some authors have reported a mean volumetric shrinkage of 11.4% and 12.4%, respectively.[22,23] This is in good agreement with the 12.4% difference we observed between volumes measured by MRI and histology. Prostate volume IOV was generally high, especially for US imaging. The border between dorsal prostate and seminal vesicles was least clear on US imaging (Figure S3) due to low contrast between the dorsal prostate and seminal vesicles. In addition, the bladder and preputial glands occasionally affected image contrast due to posterior acoustic enhancement. These factors make the transition from dorsal prostate to seminal vesicle a complex area where the placement of ROIs can be difficult and hard to reproduce. Our results indicate that the superior parts of the dorsal prostate may have been mistaken for seminal vesicles and thus were not included in the prostate ROI. While US imaging is relatively fast and cost-effective, the high IOV makes it less reliable for repeatable volume measurements than MRI. Therefore, a trade-off has to be made between the time-cost and reliability of measurements.

5 | CONCLUSIONS

US imaging was found to be a useful screening method for detecting PD tumors in the TRAMP model and for estimating PD tumor volume. MR imaging has better repeatability than US, especially for the estimation of prostate volumes and may be more suitable as a treatment monitoring method than US imaging.

ACKNOWLEDGMENTS

Animals were housed at the comparative medicine core facility, magnetic resonance imaging data were acquired at the MR core facility and histology was performed at the cellular and molecular imaging (CMIC) core facility, all at NTNU. The project is supported by the Research

Council of Norway (Project number: 240316), the Norwegian Cancer Society (Project number: 6824920-20156824920-2015), The liaison committee between the Central Norway Regional Health Authority and NTNU; (Project number: 90059700).

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

ORCID

Stein-Martin T. Fagerland http://orcid.org/0000-0002-1159-5922
Deborah K. Hill http://orcid.org/0000-0002-6441-5489
Annemieke van Wamel http://orcid.org/0000-0002-4704-2203
Catharina de Lange Davies http://orcid.org/0000-0003-4371-5025
Jana Kim http://orcid.org/0000-0003-4386-440X

REFERENCES

1. Ferlay J, Colombet M, Soerjomataram I, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries and 25 major cancers in 2018. *Eur J Cancer*. 2018;103:356-387.
2. Gengenbacher N, Singhal M, Augustin HG. Preclinical mouse solid tumour models: status quo, challenges and perspectives. *Nat Rev Cancer*. 2017;17:751-765.
3. Ittmann M, Huang J, Radaelli E, et al. Animal models of human prostate cancer: the consensus report of the New York meeting of the mouse models of Human Cancers Consortium Prostate Pathology Committee. *Cancer Res*. 2013;73:2718-2736.
4. Kido LA, de Almeida Lamas C, Marostica MR Jr, Cagnon VHA. Transgenic adenocarcinoma of the mouse prostate (TRAMP) model: a good alternative to study PCa progression and chemoprevention approaches. *Life Sci*. 2019;217:141-147.
5. Gingrich JR, Barrios RJ, Morton RA, et al. Metastatic prostate cancer in a transgenic mouse. *Cancer Res*. 1996;56:4096-4102.
6. Kaplan-Lefko PJ, Chen TM, Ittmann MM, et al. Pathobiology of autochthonous prostate cancer in a pre-clinical transgenic mouse model. *Prostate*. 2003;55:219-237.
7. Berman-Booty LD, Sargeant AM, Rosol TJ, et al. A review of the existing grading schemes and a proposal for a modified grading scheme for prostatic lesions in TRAMP mice. *Toxicol Pathol*. 2012;40:5-17.
8. Hill DK, Kim E, Teruel JR, et al. Diffusion-weighted MRI for early detection and characterization of prostate cancer in the transgenic adenocarcinoma of the mouse prostate model. *J Magn Reson Imaging*. 2016;43:1207-1217.
9. Degrassi A, Russo M, Scanziani E, et al. Magnetic resonance imaging and histopathological characterization of prostate tumors in TRAMP mice as model for pre-clinical trials. *Prostate*. 2007;67:396-404.
10. Gingrich JR, Barrios RJ, Foster BA, Greenberg NM. Pathologic progression of autochthonous prostate cancer in the TRAMP model. *Prostate Cancer Prostatic Dis*. 1999;2:70-75.
11. Suttie A, Nyska A, Haseman JK, Moser GJ, Hackett TR, Goldsworthy TL. A grading scheme for the assessment of proliferative lesions of the mouse prostate in the TRAMP model. *Toxicol Pathol*. 2003;31:31-38.
12. Shappell SB, Thomas GV, Roberts RL, et al. Prostate pathology of genetically engineered mice: definitions and classification. The consensus report from the Bar Harbor meeting of the Mouse Models of Human Cancer Consortium Prostate Pathology Committee. *Cancer Res*. 2004;64:2270-2305.
13. Chiaverotti T, Couto SS, Donjacour A, et al. Dissociation of epithelial and neuroendocrine carcinoma lineages in the transgenic adenocarcinoma of mouse prostate model of prostate cancer. *Am J Pathol*. 2008;172:236-246.
14. Søgaard CK, Moestue SA, Rye MB, et al. APIM-peptide targeting PCNA improves the efficacy of docetaxel treatment in the TRAMP mouse model of prostate cancer. *Oncotarget*. 2018;9:11752-11766.
15. Arbab AS, Shankar A, Varma NR, et al. MRI to assess chemoprevention in transgenic adenocarcinoma of mouse prostate (TRAMP). *BMC Med Imaging*. 2011;11:21.
16. Oliveira DS, Dzinic S, Bonfil AI, Saliganan AD, Sheng S, Bonfil RD. The mouse prostate: a basic anatomical and histological guideline. *Bosn J Basic Med Sci*. 2016;16:8-13.
17. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 1986;327:307-310.
18. Giavarina D. Understanding Bland Altman analysis. *Biochem Med*. 2015;25:141-151.
19. Greenberg NM, DeMayo F, Finegold MJ, et al. Prostate cancer in a transgenic mouse. *Proc Natl Acad Sci*. 1995;92:3439-3443.
20. Wirtzfeld LA, Wu G, Bygrave M, et al. A new three-dimensional ultrasound microimaging technology for preclinical studies using a transgenic prostate cancer mouse model. *Cancer Res*. 2005;65:6337-6345.
21. Linxweiler J, Körbel C, Müller A, et al. A novel mouse model of human prostate cancer to study intraprostatic tumor growth and the development of lymph node metastases. *Prostate*. 2018;78:664-675.
22. Tran T, Sundaram CP, Bahler CD, et al. Correcting the shrinkage effects of formalin fixation and tissue processing for renal tumors: toward standardization of pathological reporting of tumor size. *J Cancer*. 2015;6:759-766.
23. Schned AR, Wheeler KJ, Hodorowski CA, et al. Tissue-shrinkage correction factor in the calculation of prostate cancer volume. *Am J Surg Pathol*. 1996;20:1501-1506.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Fagerland S-MT, Hill DK, van Wamel A, de Lange Davies C, Kim J. Ultrasound and magnetic resonance imaging for group stratification and treatment monitoring in the transgenic adenocarcinoma of the mouse prostate model. *The Prostate*. 2019;1–12. https://doi.org/10.1002/pros.23930