

# Comparison between electromagnetic transponders and radiographic imaging for prostate localization: A pelvic phantom study with rotations and translations

Daniel G. Hamilton<sup>1</sup> | Dean P. McKenzie<sup>2,3</sup> | Anne E. Perkins<sup>1</sup>

<sup>1</sup>Epworth Radiation Oncology, Epworth Hospital, Richmond, Victoria, Australia

<sup>2</sup>Research, Development and Governance, Epworth Healthcare, Melbourne, Australia

<sup>3</sup>School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

Author to whom correspondence should be addressed. Daniel G. Hamilton

E-mail: danielh\_00@hotmail.com; Telephone: +61 3 9936 8277

## Abstract

The aim of this study was to evaluate the differences in target localization between Calypso<sup>®</sup>, kV orthogonal imaging and cone-beam computed tomography (CBCT) for combined translations and rotations of an anthropomorphic pelvic phantom. The phantom was localized using all three systems in 50 different positions, with applied translational and rotational offsets randomly sampled from representative normal distributions of prostate motion. Lin's concordance correlation coefficient ( $\rho_c$ ) and 95% confidence intervals were calculated to assess the agreement between the localization systems. Mean differences and difference vectors between the three systems were also calculated. Agreement between systems for lateral, vertical, and longitudinal translations was excellent, with  $\rho_c$  values of greater than 0.98 between all three systems in all axes. There was excellent agreement between the systems for rotations around the lateral axis (pitch) ( $\rho_c > 0.99$ ), and around the vertical axis (yaw) ( $\rho_c > 0.97$ ). However, somewhat poorer agreement for rotations around the longitudinal axis (roll) was observed, with the lowest correlation observed between Calypso and kV orthogonal imaging ( $\rho_c = 0.895$ ). Mean differences between the phantom position reported by Calypso and the radiographic systems were less than 1 mm and 1° for all translations and rotations. The results for translations are consistent with the publications of previous authors. There is no comparable published data for rotations. While there is lower correlation between the three systems for roll than for the other angles, the mean differences in reported rotations are not clinically significant.

## PACS

87.50.st, 87.55.km

## KEY WORDS

calypso, cone-beam computed tomography, electromagnetic transponders, localization, prostatic neoplasms, radiographic imaging, radiotherapy

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. *Journal of Applied Clinical Medical Physics* published by Wiley Periodicals, Inc. on behalf of American Association of Physicists in Medicine.

## 1 | INTRODUCTION

One of the greatest difficulties in the provision of external beam radiation therapy treatment is correcting for tumor and organ motion to ensure accurate dose delivery. For prostate radiotherapy, previous studies have shown that the shape and position of the target varies from day to day (interfraction motion) and during treatment (intrafraction motion), due to variability in patient setup, bladder and bowel filling, and patient respiration.<sup>1,2</sup> The prostate gland is liable to nontrivial combined intra- and interfractional movements and rotations, with most motion occurring in the anteroposterior and superior-inferior planes, and around the left-right axis (pitch).<sup>1-5</sup>

In current radiotherapy practice, various methods of target localization are used to correct for prostate motion.<sup>6</sup> Standard practice in Australia is to use radiographic imaging to visualize radiopaque fiducial markers implanted into the prostate.<sup>2,7</sup> The most commonly used imaging modalities are orthogonal kilovoltage planar x-rays (kV-imaging) and kilovoltage cone-beam computed tomography (CBCT). Images are usually taken prior to treatment delivery, and hence correct for interfraction motion but not intrafraction.

An alternative localization method is to use electromagnetic transponders implanted in the prostate, which can be monitored via low frequency radio waves instead of ionizing radiation. Systems using electromagnetic transponders, such as the Calypso<sup>®</sup> 4D Localization System, are being increasingly used to correct for both inter- and intrafraction prostate motion using target localization and real-time tracking of implanted transponders.<sup>8,9</sup>

Balter et al.<sup>10</sup> evaluated the accuracy and precision of this electromagnetic transponder system for translational offsets using a precisely machined mechanical jig to control the transponder position. They found that both accuracy and precision decreased with increasing distance between the transponders and the detector array, but both were less than 1 mm in all three axes over the range of geometries tested.

Other authors have compared transponder-based systems with radiographic imaging in phantom studies.<sup>11-14</sup> Santanam et al. (2009)<sup>12</sup> compared Calypso with kV orthogonal imaging in a phantom, finding agreement within 1 mm in all axes. Ogunleye et al.<sup>13</sup> also compared Calypso with kV orthogonal imaging in a phantom study, and reported submillimeter agreement between the two systems.

Patient and animal studies have also demonstrated strong correlation between transponder-based systems and radiographic systems.<sup>5,11,13,15-18</sup> Foster et al.<sup>15</sup> compared Calypso with CBCT and kV orthogonal imaging over 900 and 250 fractions respectively, and found the mean differences in localization were less than 1 mm in all axes. Ogunleye et al.<sup>13</sup> extended their phantom study to a cohort of 259 patient measurements, again demonstrating good correlation between kV orthogonal imaging and Calypso. The results of Wiloughby et al.<sup>16</sup> were similar, with a mean 3D distance vector difference of 1.5 mm between Calypso and orthogonal kV-imaging. Quigley et al.<sup>17</sup> compared Calypso with radiographic monitoring over 1027 fractions using the ExacTrac system and found that the mean

vector length difference between Calypso and ExacTrac was 1.9 mm  $\pm$  1.2 mm. Commissioning tests performed in our own center showed similar agreement between Calypso and radiographic imaging for translations, with mean 3D distance vector differences of 1.1 mm and 1.5 mm for Calypso-kV orthogonal and Calypso-CBCT, respectively.

While the studies described above have demonstrated strong correlation between transponder-based systems and radiographic systems for translational offsets, little has been published about the accuracy of rotational offsets. Santanam et al. (2009)<sup>12</sup> investigated the accuracy of Calypso rotation as part of their commissioning procedure. They tilted a phantom through 20° of pitch and roll by positioning it on a foam wedge, and compared the rotational values reported by Calypso with those recorded using a digital level. They tested yaw by rotating the treatment couch. The authors' reported agreement within 1° for all rotations, but only tested five positions in total (two each for pitch and roll, one for yaw). Commissioning tests conducted in our own center using similar methods showed similar results. Li et al. (2009)<sup>19</sup> also checked the rotational accuracy of Calypso in a phantom and reported that it was within 1°, but did not provide details of their methods or results.

In spite of this paucity of evidence evaluating the accuracy of transponder-based systems like Calypso in measuring rotational offsets, the authors note that a number of recent studies have used Calypso-generated localization data to draw conclusions on margin calculations and dosimetric coverage.<sup>20-23</sup> For example, the results of one study using this rotational offset information suggested that inter- and intrafraction prostatic rotations may result in target underdosing in up to 61% of patients depending on the PTV expansion used.<sup>20</sup> The same study above also proposed that this information may be useful in the future in developing a metric to predict target coverage in the clinic prior to radiotherapy treatment.

In consideration of the current and future usage of Calypso-derived rotational offset information, the aim of this phantom study was therefore to determine the level of agreement between Calypso with both kV orthogonal imaging and cone-beam CT for a series of realistic combined rotations and translations, representative of typical prostate motion as described in the literature.<sup>3,24</sup>

## 2 | METHODS AND MATERIALS

### 2.A | Study design

This was an observational quality assurance (QA) study to assess the accuracy of the Calypso<sup>®</sup> 4D localization system compared with CBCT and kV orthogonal imaging. Numerous statistical analyses are available to determine the level of agreement between two methods of measuring the same continuous variable. Conventional measures of correlation, such as Pearson's, can achieve their maximum value of +1/-1 even when there is no agreement between measures (i.e., the Pearson or Spearman correlation between scores 1,2,3 and 4,5,6 is 1, yet no pair of scores is equal). Other measures, such as the intraclass correlation, and the more recent Lin's concordance

coefficient ( $\rho_c$ ),<sup>25,26</sup> only achieve a value of 1 when two sets of scores are identical to each other.

In line with McBride's<sup>27</sup> recommendations, Lin's concordance correlation was used to evaluate agreement between the localization systems for this study. As such, based on the results of a previous study,<sup>15</sup> a concordance correlation coefficient of at least 0.80 between each pair of techniques was expected. As sample size calculators based directly upon the expected confidence interval for a concordance correlation coefficient are not widely available, sample sizes were based upon the broadly similar intraclass correlation coefficient, with an expected value of 0.80, and 80% assurance of obtaining a 95% confidence.<sup>28</sup>

Previously published studies of internal prostate movement confirmed that rotations and translations follow approximately normal distributions, and provided values for the associated means and standard deviations.<sup>3,24</sup> Three sets each of 50 translational and 50 rotational offsets were randomly sampled from normal distributions generated using the aforementioned specified means and standard deviations from references<sup>3,24</sup> using Monte Carlo methods, obtained using the RANNOR function of SAS 9.4 (SAS Institute Incorporated, Cary, North Carolina, 2013). These 50 sets of values were used to generate the table of phantom translations and rotations used for the study (Table S1).

## 2.B | Simulation and planning

The Calypso system has previously been described in detail.<sup>10,17</sup> For this study, three Calypso beacon transponders were inserted into a radiolucent foam cylinder representing the prostate gland. The transponders were positioned in the shape of an equilateral triangle with side length approximately 3 cm, in accordance with the manufacturer's recommendations. The foam cylinder was subsequently placed inside an anthropomorphic pelvic phantom (CIRS pelvic phantom) (Fig. 1). The phantom was scanned on a Siemens SOMATOM Emotion CT scanner (Siemens Medical Systems, Forchheim, Germany) with 1 mm slices.



**FIG. 1.** CIRS pelvic phantom.

Images were transferred to Varian's Eclipse™ treatment planning system (Version 11.0.42) (Varian Medical Systems, Palo Alto, California). Following standard planning procedures for Calypso patients, a treatment plan was created with the isocenter at the center of mass of the transponders. CBCT and orthogonal kV setup fields were added to the plan, to allow radiographic acquisition at the linear accelerator. Standard departmental procedures were used to transfer the treatment plan and associated imaging fields to Varian's Aria record and verify system (Version 13.6) and the Calypso tracking station (Version 3.0).

## 2.C | Phantom positioning, localization, and image acquisition

The pelvic phantom was positioned and imaged 50 times for this study according to the set of pregenerated rotational and translational offsets (Table S1). Prior to localization, the phantom was placed on a board on the treatment couch top. The desired yaw was achieved by rotating the phantom on the board about the collimator rotation axis, using an angular scale marked on the phantom with the sagittal laser as a reference. The desired pitch and roll were achieved by tilting the board by placing spacers underneath it. Pitch and roll were measured using a 2-axis digital level (Digi-Pas DWL3000xy, Digipas Technologies Inc., Irvine, California). The position of the phantom on the board and the tilt of the board were iteratively adjusted until pitch, yaw, and roll were each within 0.5° of the values specified in the table.

After applying the desired angular offsets, the phantom was positioned at the isocenter according to Calypso, as is normal clinical practice. While any localization system could have been used for this initial positioning, Calypso was used because it was the quickest method. The desired translational offsets were then applied by moving the treatment couch in the lateral, vertical, and longitudinal directions. The setup procedure resulted in the entire phantom being rotated and translated, rather than an internal rotation and translation of the prostate within the phantom. This was done mainly for measurement efficiency, since moving the prostate internally would have required dismantling and re-assembling the phantom for each measurement. Moreover, it would not be possible to set internal yaw and pitch rotations since the foam "prostate" must fit within a machined cylindrical cavity inside the phantom.

Once the phantom was in the required location, the transponders were localized according to Calypso, with offsets recorded from the console screen. These values were subsequently verified by the associated session reports. Radiographic images were then acquired using the On-Board Imager (OBI) (Version 1.6) and saved for later analysis. The images consisted of an orthogonal kV pair (right lateral and PA), and a full-fan CBCT using the "Pelvis spotlight" protocol. This protocol uses a small field of view which provides high resolution in the central pelvis at the cost of missing some of the skin surface. All data were acquired on the same linear accelerator, a Varian Trilogy™. The 50 measurements required for the study were

performed on three separate days over a twelve-week period, with four measurements performed on the first day, 32 on the second and 14 on the last day.

## 2.D | Radiographic image analysis

The kV orthogonal images and CBCT images were analyzed offline using the Varian™ Offline Review software (Version 13.6) (Fig. 2). For the purposes of the study, the software was configured to report rotational offsets about all three axes (pitch, roll, and yaw), although in our normal clinical practice only yaw is reported.

For the kV orthogonal images, the Marker Match option was used (Fig. 2). This matching algorithm allows the user to mark the position of fiducial markers in a pair of orthogonal images. The software then calculates translational and rotational offsets in three dimensions for the image pair. The centers of the three Calypso transponders were used as the fiducial markers. For the CBCT, the automated anatomy matching software was used. All the anatomy including bone and soft tissue analogs was used. The field of view for matching was set to encompass just the CBCT image, approximately 2 cm inside the phantom surface. The automated anatomy matching algorithm uses a mutual information algorithm to align the CBCT with the planning CT data set, again reporting translational and rotational offsets in three dimensions. Both of these matching algorithms were chosen specifically in order to minimize the operator-dependent uncertainty in matching, which would be greater if manual matching was used.

All images were acquired, matched and entered by one author (AEP) into a preprepared spreadsheet, then checked for accuracy by the primary author (DGH). As Calypso and Aria use different conventions for reporting rotations and translations, all translation and rotation values from the radiographic imaging systems were corrected to match the Calypso conventions, which are as described in Santanam et al. (2009).<sup>12</sup>

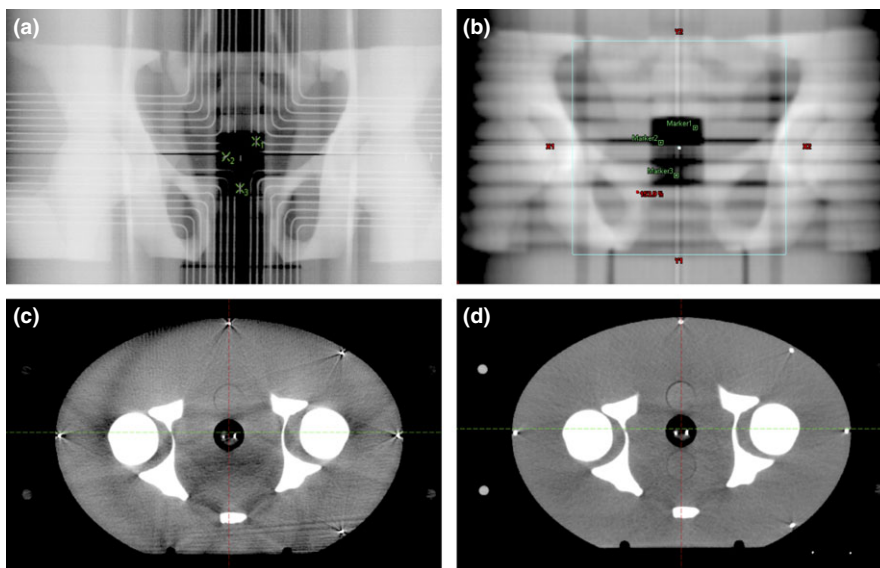
## 2.E | Statistical analysis

Lin's concordance correlation coefficients ( $\rho_c$ ) and 95% confidence intervals were calculated to assess the degree of correlation between Calypso-kV, Calypso-CBCT, and kV-CBCT localizations. The differences between Calypso-kV, Calypso-CBCT, and kV-CBCT were calculated for all three translational axes (lateral, vertical, and longitudinal) and all three rotations (pitch, roll, and yaw). The mean difference, standard deviation, and range for the 50 measurements were then derived. The vector length differences between the three imaging systems were also calculated.<sup>13</sup> Offset values were generated using SAS (Version 9.4, SAS Institute Incorporated, Cary, North Carolina, 2013) as described above. All other analyses were performed using Stata (version 14) (Stata Corporation, College Station, Texas, 2015).

As described above, Lin's concordance correlation coefficient is used to determine the level of agreement between two methods of measuring the same continuous variable. Different authors have made suggestions for interpreting the scores for Lin's concordance correlation coefficients.<sup>27,29</sup> One such proposal by McBride,<sup>27</sup> in assessing the degree of equivalence between new laboratory tests and the gold standard, has suggested that a lower 95% confidence interval greater than 0.99 represents almost perfect agreement, 0.95 to 0.99 substantial agreement and 0.90 to 0.95 moderate agreement. Although McBride<sup>27</sup> was considering a different specific application (laboratory tests in microbiology), his criteria appear to be appropriate for this study.

## 2.F | Quality assurance checks of localization systems

During the period of the study, the coincidence between the lasers, Calypso, the kV-imaging system and the MV isocenter was checked according to standard departmental monthly quality assurance procedures. These are based on the recommendations of AAPM Task Group 142<sup>30</sup> for a linear accelerator used for IMRT and SRS/SBRT.



**FIG. 2.** (a) Anterior kV image showing the Calypso transponders (green crosses) and copper wires in electromagnetic array, (b) planning DRR with markers indicating Calypso transponders, (c) CBCT, and (d) planning CT scan.



The Task Group 142 report does not give recommendations for QA of electromagnetic transponder systems. Because Calypso is an alternative localization system, we apply the same tests and tolerances to Calypso as to the kV-imaging systems. Briefly, the QA procedure consists of locating the treatment radiation isocenter via a Winston-Lutz test, and then confirming that all the localization systems (lasers, kV planar imaging, CBCT and Calypso) indicate the radiation isocenter within 1 mm in any axis. All QA tests were within tolerance and so no adjustments were made to any of the localization systems during the course of the study.

### 3 | RESULTS

#### 3.A | Translations

Scatter plots showing the translational offsets reported by Calypso and the radiographic imaging systems are shown in Fig. 3. Also shown in Fig. 3 are corresponding linear regression lines and the line of identity. The line of identity is the line along which all data would fall if there was perfect agreement between the systems. The linear regression line and line of identity are virtually indistinguishable in all the plots except for the lateral axis where there is an offset of approximately 0.6 mm between Calypso and the radiographic systems. Error bars are not shown on the graphs for reasons of clarity. Please refer to the appendix for a discussion of the uncertainties associated with the data in this figure.

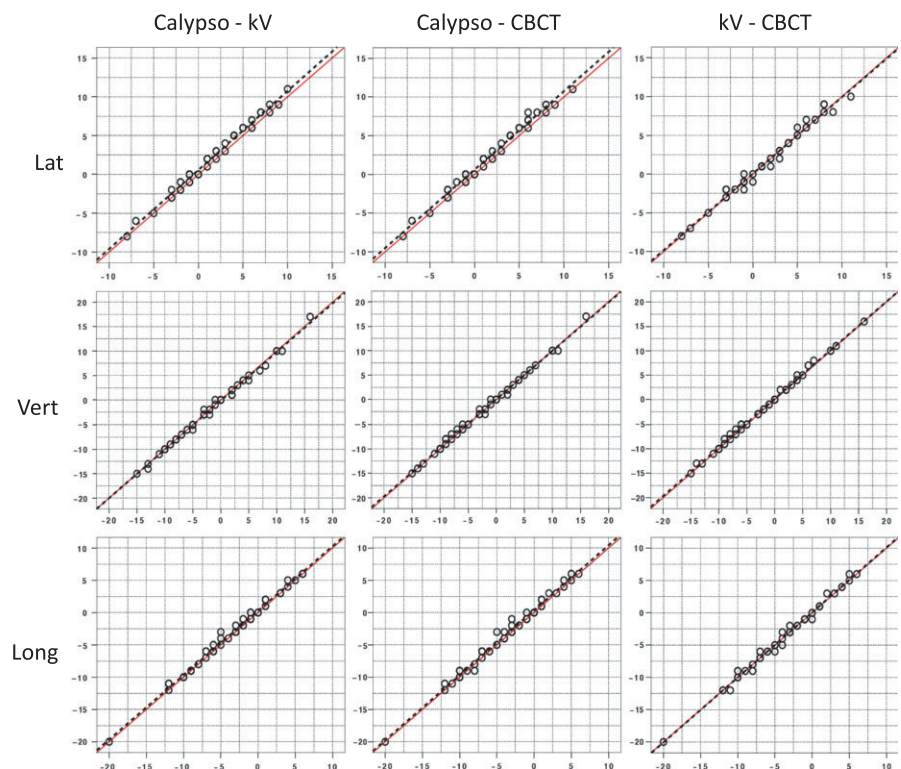
Lin's concordance correlation coefficients ( $\rho_c$ ) between the three systems for translational localizations, with the associated 95% confidence intervals, are displayed in the upper half of Table 1.

Agreement between all three systems in the vertical and longitudinal directions is excellent, with observed  $\rho_c$  values of greater than 0.990 for all comparisons. In the lateral direction, there is also excellent correlation between kV orthogonal and CBCT ( $\rho_c = 0.992$ , 95% CI 0.986–0.995). However, there is lower agreement between Calypso-kV orthogonal ( $\rho_c = 0.983$ , 95% CI 0.972–0.990) and Calypso-CBCT ( $\rho_c = 0.981$ , 95% CI 0.969–0.988).

Differences in translations between the three imaging systems are summarized in the upper half of Table 2, which shows the mean, standard deviation, and range for translational differences in the lateral, vertical, and longitudinal axes over the 50 measurements. The mean, standard deviation, and range for the difference vector length are also shown. The mean differences are all less than 1 mm in all three axes, as are the upper and lower 95% confidence intervals (results not reported). The maximum difference observed in any axis was 2 mm. The mean difference vector length is 1 mm or less in all cases, and the maximum length is 2.5 mm.

#### 3.B | Rotations

Scatter plots showing the rotational offsets reported by Calypso and the radiographic imaging systems are shown in Fig. 4, with the linear regression line and the line of identity. The two lines are virtually indistinguishable for pitch, but for roll there is a noticeable difference in slope, particularly when comparing Calypso with the radiographic imaging systems. There is also a difference in slope of the two lines for yaw when comparing Calypso with kV orthogonal. Error bars are not shown in Fig. 4 for clarity, but sources of uncertainty are discussed in the appendix.



**FIG. 3.** Scatter plots comparing all three localization systems for 50 measured lateral (top), vertical (middle), and longitudinal (bottom) translations. (Red lines refer to the line of identity and dashed black lines to the corresponding linear regression line.)

**TABLE 1** Lin's concordance correlation coefficient ( $\rho_c$ ) and 95% confidence intervals for localization values between Calypso-kV imaging, Calypso-CBCT and kV imaging-CBCT for selected translations (mm) and rotations (deg).

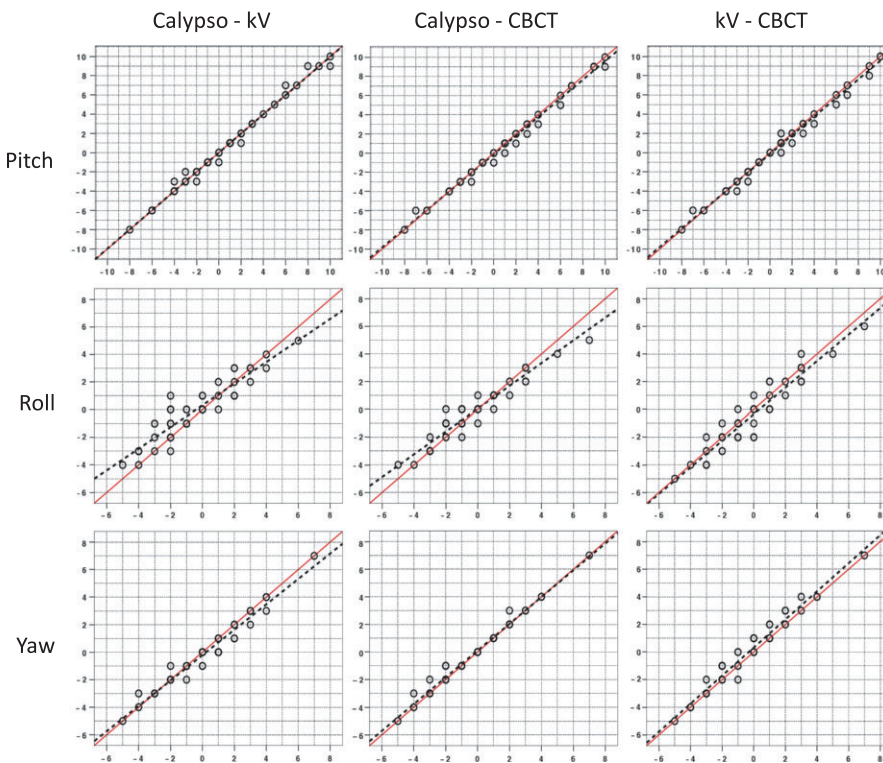
|              | Calypso-kV imaging |             | Calypso-CBCT |             | kV imaging-CBCT |             |
|--------------|--------------------|-------------|--------------|-------------|-----------------|-------------|
|              | $\rho_c$           | 95% CI      | $\rho_c$     | 95% CI      | $\rho_c$        | 95% CI      |
| Lateral      | 0.983              | 0.972–0.990 | 0.981        | 0.969–0.988 | 0.992           | 0.986–0.995 |
| Vertical     | 0.997              | 0.995–0.999 | 0.998        | 0.996–0.999 | 0.998           | 0.996–0.999 |
| Longitudinal | 0.994              | 0.990–0.997 | 0.992        | 0.986–0.995 | 0.996           | 0.993–0.998 |
| Pitch        | 0.995              | 0.992–0.997 | 0.995        | 0.991–0.997 | 0.994           | 0.989–0.996 |
| Roll         | 0.895              | 0.830–0.936 | 0.947        | 0.914–0.967 | 0.939           | 0.896–0.965 |
| Yaw          | 0.975              | 0.957–0.985 | 0.989        | 0.980–0.993 | 0.968           | 0.946–0.981 |

CBCT, cone-beam computed tomography; kV, kilovoltage; CI, confidence interval.

**TABLE 2** Mean differences in localization between Calypso-kV imaging, Calypso-CBCT and kV imaging-CBCT for random translations (mm) and rotations (deg).

|                 | Calypso-kV imaging |     |          | Calypso-CBCT |     |          | kV imaging-CBCT |     |          |
|-----------------|--------------------|-----|----------|--------------|-----|----------|-----------------|-----|----------|
|                 | Mean               | SD  | Range    | Mean         | SD  | Range    | Mean            | SD  | Range    |
| Lateral         | 0.6                | 0.5 | (0, 1)   | 0.6          | 0.5 | (0, 2)   | 0.0             | 0.5 | (-1, 1)  |
| Vertical        | -0.1               | 0.5 | (-1, 1)  | 0.1          | 0.5 | (-1, 1)  | 0.2             | 0.4 | (0, 1)   |
| Longitudinal    | 0.3                | 0.5 | (0, 2)   | 0.3          | 0.6 | (-1, 2)  | 0.0             | 0.5 | (-1, 1)  |
| $\Delta$ Vector | 0.9                | 0.5 | (0, 2.2) | 1.0          | 0.6 | (0, 2.5) | 0.6             | 0.6 | (0, 1.4) |
| Pitch           | 0.0                | 0.4 | (-1, 1)  | -0.1         | 0.4 | (-1, 1)  | -0.1            | 0.5 | (-1, 1)  |
| Roll            | 0.4                | 1.0 | (-1, 3)  | 0.1          | 0.7 | (-2, 2)  | -0.3            | 0.8 | (-2, 1)  |
| Yaw             | -0.2               | 0.5 | (-1, 1)  | 0.1          | 0.3 | (0, 1)   | 0.3             | 0.5 | (-1, 1)  |

CBCT, cone-beam computed tomography; kV, kilovoltage; SD, standard deviation;  $\Delta$ Vector, difference vector magnitude.



**FIG. 4.** Scatter plots comparing all three localization systems for 50 measured pitch (top), roll (middle), and yaw (bottom) rotations. (Red lines refer to the line of identity and dashed black lines the corresponding linear regression line.)

Concordance correlation coefficients ( $\rho_c$ ) between all three systems for rotational localizations are displayed in the lower half of Table 1. Agreement between all three systems for pitch was excellent, with  $\rho_c$  values greater than 0.990 in all cases. For yaw, the

degree of correlation is less, with  $\rho_c$  values in the range 0.968–0.989. The degree of correlation with roll is noticeably poorer, with  $\rho_c$  values of 0.895 for Calypso-kV imaging, 0.947 for Calypso-CBCT, and 0.939 for kV imaging-CBCT.

Differences in rotations between the three imaging systems are summarized in the lower half of Table 2, which shows the mean, standard deviation, and range for differences in pitch, roll, and yaw. The mean differences are less than 1° in all three cases. The maximum difference observed for any rotation was 3°, which was recorded when comparing Calypso with kV orthogonal imaging for roll.

## 4 | DISCUSSION

The aim of this phantom study was to prospectively compare the accuracy of Calypso with kV orthogonal imaging and CBCT in detecting combined translations and rotations over a range of offsets representative of typical prostate motion. Unlike previously published studies, the current work has investigated rotational accuracy in detail, and in combination with translations.

### 4.A | Translations

Our results for translations are similar to those published by other authors in phantom and patient studies.<sup>13,15,16,18</sup> Table 3 shows our mean differences and standard deviations for translations in the lateral, vertical and longitudinal directions, and for the vector translation, in comparison with the results of these studies. The results of all studies are in agreement within two standard deviations. The highest mean difference in any axis reported by any study is 1.2 mm,<sup>13</sup> and all other studies show submillimeter agreement.

Although there is good agreement between all the results in Table 3, the data suggest that some institutions achieve better mean agreement between Calypso and their radiographic imaging systems than others, and that agreement is better in some directions than others. This is supported by our own data, which shows evidence of a small offset in the lateral direction between Calypso and the radiographic imaging systems (see Fig. 3). We believe that this represents a small systematic error of approximately 0.6 mm in the calibration

of our localization systems, although well within our monthly QA tolerance of 1 mm.

While these results are encouraging, the data presented in this report only enable us to identify a discrepancy between Calypso and the radiographic systems — it does not tell us how the systems compare to the megavoltage radiation isocenter, which is the fundamental reference for calibration of localization systems. This could be determined by performing the monthly quality assurance check described in the Methods section. When multiple localization systems are available in a department, there may be a need to consider tighter tolerances on each individual system than specified in the AAPM TG142 report.<sup>30</sup> In the extreme case, two localization systems may each be within 1 mm of the megavoltage radiation isocenter and hence within the AAPM TG142 tolerance, but may differ by 2 mm from each other. Such disagreement can pose clinical decision-making problems when systems are used together for patient localization, particularly when tight treatment margins are used.

Table 3 also shows that the standard deviations are generally lower in the phantom studies than in patient studies. This is not surprising. In the patient studies, intrafraction motion secondary to respiration and bladder and bowel filling will contribute to differences in transponder position if the radiographic imaging and Calypso monitoring are performed at different times. Furthermore, the algorithms used by Calypso and the radiographic systems to calculate patient position may handle rotations and distortions differently, contributing to increased uncertainty in the reported positions.

### 4.B | Rotations

Our data on rotations are, so far as we are aware, the first published which systematically compares the rotational accuracy of Calypso with radiographic imaging over a range of angular offsets. We have found overall excellent agreement between Calypso and the radiographic systems for pitch, roll, and yaw over a range of clinically relevant angles. The mean differences between each system are less than 1°, and the largest discrepancy was 3°. These data are consistent with the results reported by Santanam et al. (2009)<sup>12</sup> and Li

**TABLE 3** Previously reported mean differences in localization (magnitude only) between Calypso and radiographic imaging for translations (mm).

| Author                   | Study type | N    | Radiographic comparator | Mean difference (SD) |           |           |           |
|--------------------------|------------|------|-------------------------|----------------------|-----------|-----------|-----------|
|                          |            |      |                         | Lat                  | Vert      | Long      | Vector    |
| Current study            | Phantom    | 50   | kV-imaging              | 0.6 (0.5)            | 0.1 (0.5) | 0.3 (0.5) | 1.0 (0.5) |
|                          | Phantom    | 50   | kV-CBCT                 | 0.6 (0.5)            | 0.1 (0.5) | 0.3 (0.6) | 1.0 (0.6) |
| Willoughby <sup>16</sup> | Phantom    | NR   | kV-imaging              | –                    | –         | –         | 0.5 (0.1) |
|                          | Patient    | 11   | kV-imaging              | –                    | –         | –         | 1.5 (0.9) |
| Ogunleye <sup>13</sup>   | Phantom    | 30   | kV-imaging              | 0.4 (0.4)            | 0.2 (0.3) | 0.4 (0.3) | 0.8 (0.4) |
|                          | Patient    | 259  | kV-imaging              | 0.7 (0.5)            | 1.2 (0.9) | 1.1 (0.9) | 2.1 (1.0) |
| Kupelian <sup>18</sup>   | Patient    | 1027 | ExacTrac                | 0.1 (0.9)            | 0.0 (1.3) | 0.4 (1.4) | 1.9 (1.2) |
| Foster <sup>15</sup>     | Patient    | 260  | kV-imaging              | 0.1 (1.0)            | 0.1 (1.4) | 0.5 (1.5) | –         |
|                          | Patient    | 915  | kV-CBCT                 | 0.0 (1.2)            | 0.2 (2.9) | 0.8 (2.2) | –         |

SD, standard deviation; kV, kilovoltage; CBCT, cone-beam computed tomography.

et al. (2009).<sup>19</sup> However, the Lin's concordance correlation coefficient results show that agreement between the systems is best for pitch, somewhat worse for yaw, and worse again for roll. The biggest disagreement was found between Calypso and kV orthogonal imaging for determination of roll ( $\rho_c = 0.895$ , see Table 1). It is difficult to be certain of the reasons for this without full knowledge of the algorithms used by all the systems to determine rotations. We believe that one factor is a fundamental limitation of matching using orthogonal kV images. Roll is a rotation around the longitudinal axis, which is also the axis around which the gantry rotates when taking orthogonal kV images. Visual inspection of kV orthogonal images shows that yaw (rotation about the vertical axis) is readily apparent in a PA image, and pitch (rotation about the lateral axis) is readily apparent in a lateral image, but roll is only indirectly visible as changes in the apparent separation of the fiducial markers in both the AP and lateral images. Even without knowing the details of the marker match algorithm, it is not surprising that roll may be partially misinterpreted as translation when using orthogonal image pairs. This suspicion is given additional weight by our uncertainty analysis, which shows a greater uncertainty in the kV orthogonal matching for roll than for pitch and yaw (see appendix).

Detection of roll in CBCT images should not be affected in the same way as kV orthogonal pairs, since the matching algorithm uses the full 3D data set. This is also confirmed by our uncertainty analysis (see appendix). Further confirmation comes from inspection of our raw experimental data (not presented here) which shows that CBCT gave results that matched our digital level readings within 1° for all roll settings, whereas both Calypso and kV orthogonal matching differed from the digital level by up to 2°.

#### 4.C | Lin's concordance correlation coefficient

Applying the criteria described by McBride<sup>27</sup> to our results indicates that we have "almost perfect agreement" between the three localization systems for vertical and longitudinal translations. This is in agreement with the data shown in Fig. 3 and Table 2, which confirms that the systems give essentially equivalent results. McBride's<sup>27</sup> terminology of "almost perfect agreement" could be interpreted in the radiotherapy context as meaning that a patient could be positioned with any of the systems with the same level of accuracy.

For lateral translations, our data comparing Calypso to the radiographic systems would fall into the "substantial agreement" category. As noted previously, there appears to be a small offset of approximately 0.6 mm between Calypso and the radiographic imaging systems (see Table 2 and Fig. 3). Although the systems have an equivalent ability to detect translational offsets, there is likely, however, a systematic difference between them attributable to a different reference point being established at the time the systems were calibrated. In the context of this study, "substantial agreement" could be interpreted as "within calibration tolerances". Applying these criteria to our rotation data indicates almost perfect agreement for pitch, substantial agreement for yaw but only moderate agreement for roll.

#### 4.D | Study limitations

While this study was successful in its aim of systematically comparing Calypso with radiographic imaging for a range of clinically relevant combined rotations and translations, it does have a number of limitations.

Because the aim was to investigate the capabilities of all the systems under the best possible circumstances, we used a number of procedures that are not part of our normal clinical practice for prostate patients including:

1. Using the Marker Match algorithm for the kV orthogonal images, when manual matching is our normal clinical practice.
2. Using automatic anatomy matching for the CBCT images, when manual matching to markers is our normal clinical practice.
3. Reporting roll, pitch and yaw for the radiographic images (both kV orthogonal and CBCT), when in clinical practice only yaw is reported.

The results of our study would be expected to differ if we used our normal clinical match settings. Results may also differ if a different individual performed the matching. Further investigations are planned to test the impact of these practices.

Although we used a pelvic phantom for the study, in order to achieve a realistic patient geometry for testing the systems, our experimental setup had some limitations. One of these was that we chose to rotate and translate the whole phantom rather than moving the prostate internally. As noted in the methods section, this was done primarily for practical reasons, to enable the measurements to be carried out in a reasonable time frame without the need to disassemble the phantom for each measurement. Although rotating the whole phantom is not representative of normal patient treatment, we feel that this is was not of major significance for our phantom based study, since in a phantom-based study the relationship of all components of the phantom remains fixed.

Another potential phantom-related limitation of this study is the relatively small size of the CIRS pelvic phantom (20 cm in the anteroposterior direction by 30 cm in the lateral direction). Calypso cannot be used for men with a substantially large body habitus, as the manufacturer has built in a software restriction that limits the maximum distance between the detector array and the transponders. Investigations performed during commissioning of Calypso in our center showed that phantom size did not significantly affect the positional accuracy of Calypso over the range of allowed array distances, although random noise in the Calypso signal increased slightly for larger phantoms, as the distance between the transponders and the detector panel increased. These results are in agreement with those reported by Balter et al.<sup>10</sup> who found submillimeter accuracy was maintained as distance from the array to the transponders was increased up to the maximum allowed value of 27 cm. In theory, the accuracy of radiographic imaging would be expected to decrease slightly with phantom size due to reduced signal to noise ratio, but our clinical experience shows that this is not significant for patients eligible for Calypso treatment.



Another limitation of our experimental setup is that the Calypso beacons and isocenter were positioned as per manufacturer guidelines. In patients, such ideal placement of the beacons and isocenter is not always possible, and under such situations we would expect the level of agreement between the localization systems to worsen as the beacon geometry becomes more collinear and the separation of the beacons reduces. Positioning the isocenter away from the beacon center of mass would not affect our results however. All three systems can still detect the beacon position regardless of the isocenter location, provided that the beacons do not move outside the system's field of view.

Furthermore, we kept the beacons fixed in the same spatial relationship to each other throughout the experiment. Literature shows that in patients, fiducial markers implanted in the prostate can migrate, as well as move relative to each other.<sup>2,31</sup> Fiducial deformation is also commonplace among patients receiving postprostatectomy radiotherapy where transponders are inserted into the prostatic fossa.<sup>32</sup> No attempt was made in this study to assess the effects of deformations and transponder movement. It is expected that the algorithms used by Calypso and the radiographic imaging will handle deformations differently, and hence there are likely to be greater discrepancies between the systems in patients than in this idealized phantom geometry, as reflected in the publications of other authors summarized in Table 3.

#### 4.E | Clinical implications

The results of our study confirm the findings of others that Calypso gives a similar level of accuracy for patient translations as current best practice radiographic imaging.<sup>13,15,16</sup> In addition, we now have increased confidence in the accuracy of the rotational information provided by Calypso. This information is important for our center in establishing the reliability of Calypso, so that our patients can benefit from the real-time intrafraction motion detection. It also opens up the possibility of decreasing the use of radiographic imaging, thereby reducing patient imaging dose with negligible impact on treatment efficiency.<sup>17</sup>

In theory, when a rotation is detected, this could be corrected for by changing patient tilt position (e.g., using a 6D couch). This however is not standard practice for prostate radiotherapy, although it is commonly used for stereotactic work. The negative impact of prostate rotation on dosimetric coverage has been widely recognized.<sup>20,21,33</sup> Consequently, the discovery of large rotational offsets upon prostate localization in the clinic will impact clinical decision-making. When Calypso detects rotation outside the predefined limits, it warns the operator. In our center, this warning serves as a trigger to investigate patient set up and to perform soft tissue imaging to assess bladder and bowel filling.

To date, there have been several studies that have utilized Calypso-generated rotational offsets to evaluate PTV margins and dosimetric coverage for prostate radiotherapy.<sup>20–23</sup> These studies show that target rotations can cause significant target under-dosing even when translations are corrected for, with negative effects

complexly linked to a combination of the magnitude of rotation, target shape and the distance between the target and transponders' rotational centroid (eccentricity).<sup>20,22</sup> As such, while there is not yet a universally recognized trigger value for action, validation of the integrity of Calypso-derived rotational offset information is valuable, particularly if these values are to be used in the future as evidence to assist in clinical decision-making.

## 5 | CONCLUSION

This phantom study has compared Calypso with kV orthogonal imaging and CBCT in measuring combined translations and rotations over a range of movements representative of typical prostate motion. Unlike previously published studies, the current work has investigated rotational accuracy in detail, and in combination with translations. Our results confirm the results of other authors with regard to translations, with submillimeter differences and substantial to almost perfect agreement observed between the three systems. For rotations, varying degrees of agreement were observed depending on the rotational axis. In spite of this, mean differences were less than 1° for all three systems, which is adequate for clinical use. These results give confidence in the use of Calypso translational and rotational data for patient positioning, margin calculation, and treatment decision-making.

#### ACKNOWLEDGMENTS

The authors acknowledge and thank the Australian Clinical Dosimetry Service for provision of the CIRS phantom.

#### CONFLICTS OF INTEREST

The authors report no conflicts of interest or financial interests. The authors alone are responsible for the content and writing of the paper.

#### RESEARCH FUNDING

This work was supported by the Epworth Research Institute [grant number 11.952.000.80992].

#### REFERENCES

1. Dinkel J, Thieke C, Plathow C, et al. Respiratory-induced prostate motion: characterization and quantification in dynamic MRI. *Strahlenther Onkol*. 2011;187:426–432.
2. Ng M, Brown E, Williams A, Chao M, Lawrentschuk N, Chee R. Fiducial markers and spacers in prostate radiotherapy: current applications. *BJU Int*. 2014;113:13–20.
3. Graf R, Boehmer D, Budach V, Wust P. Interfraction rotation of the prostate as evaluated by kilovoltage X-ray fiducial marker imaging in intensity-modulated radiotherapy of localized prostate cancer. *Med Dosim*. 2012;37:396–400.
4. Langen KM, Willoughby TR, Meeks SL, et al. Observations on real-time prostate gland motion using electromagnetic tracking. *Int J Radiat Oncol Biol Phys*. 2008;71:1084–1090.

5. Mayyas E, Chetty IJ, Chetvertkov M, et al. Evaluation of multiple image-based modalities for image-guided radiation therapy (IGRT) of prostate carcinoma: a prospective study. *Med Phys*. 2013;40:041707.
6. Das S, Liu T, Jani AB, et al. Comparison of image-guided radiotherapy technologies for prostate cancer. *Am J Clin Oncol*. 2014;37:616–623.
7. Batumalai V, Holloway LC, Kumar S, et al. Survey of image-guided radiotherapy use in Australia. *J Med Imaging Radiat Oncol*. 2016;61:394–401.
8. Shah AP, Kupelian PA, Willoughby TR, Meeks SL. Expanding the use of real-time electromagnetic tracking in radiation oncology. *J Appl Clin Med Phys*. 2011;12:3590.
9. D'Ambrosio DJ, Bayouth J, Chetty IJ, et al. Continuous localization technologies for radiotherapy delivery: report of the american society for radiation oncology emerging technology committee. *Pract Radiat Oncol*. 2012;2:145–150.
10. Balter JM, Wright JN, Newell LJ, et al. Accuracy of a wireless localization system for radiotherapy. *Int J Radiat Oncol Biol Phys*. 2005;61:933–937.
11. Santanam L, Malinowski K, Hubenschmidt J, et al. Fiducial-based translational localization accuracy of electromagnetic tracking system and on-board kilovoltage imaging system. *Int J Radiat Oncol Biol Phys*. 2008;70:892–899.
12. Santanam L, Noel C, Willoughby TR, et al. Quality assurance for clinical implementation of an electromagnetic tracking system. *Med Phys*. 2009;36:3477.
13. Ogunleye T, Rossi PJ, Jani AB, Fox T, Elder E. Performance evaluation of calypso 4D localization and kilovoltage image guidance systems for interfraction motion management of prostate patients. *Sci World J*. 2009;9:449–458.
14. Rau AW, Nill S, Eidsens RS, Oelfke U. Synchronized tumour tracking with electromagnetic transponders and kV x-ray imaging: evaluation based on a thorax phantom. *Phys Med Biol*. 2008;53:3789–3805.
15. Foster RD, Pistenmaa DA, Solberg TD. A comparison of radiographic techniques and electromagnetic transponders for localization of the prostate. *Radiation Oncology*. 2012;7:101.
16. Willoughby TR, Kupelian PA, Pouliot J, et al. Target localization and real-time tracking using the Calypso 4D localization system in patients with localized prostate cancer. *Int J Radiat Oncol Biol Phys*. 2006;65:528–534.
17. Quigley MM, Mate TP, Sylvester JE. Prostate tumor alignment and continuous, real-time adaptive radiation therapy using electromagnetic fiducials: clinical and cost-utility analyses. *Urol Oncol*. 2009;27:473–482.
18. Kupelian P, Willoughby T, Mahadevan A, et al. Multi-institutional clinical experience with the calypso system in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy. *Int J Radiat Oncol Biol Phys*. 2007;67:1088–1098.
19. Li JS, Jin L, Pollack A, et al. Gains from real-time tracking of prostate motion during external beam radiation therapy. *Int J Radiat Oncol Biol Phys*. 2009;75:1613–1620.
20. Amro H, Hamstra DA, Mcshan DL, et al. The dosimetric impact of prostate rotations during electromagnetically guided external-beam radiation therapy. *Int J Radiat Oncol Biol Phys*. 2013;85:230–236.
21. Li JS, Jin L, Pollack A, et al. Gains from real-time tracking of prostate motion during external beam radiation therapy. *Int J Radiat Oncol Biol Phys*. 2009;75:1613–1620.
22. Olsen JR, Noel CE, Baker K, Santanam L, Michalski JM, Parikh PJ. Practical method of adaptive radiotherapy for prostate cancer using real-time electromagnetic tracking. *Int J Radiat Oncol Biol Phys*. 2012;82:1903–1911.
23. Wu J, Ruan D, Cho B, et al. Electromagnetic detection and real-time DMLC adaptation to target rotation during radiotherapy. *Int J Radiat Oncol Biol Phys*. 2012;82:e545–e553.
24. Chen J, Lee JL, Handrahan D, Sause WT. Intensity-modulated radiotherapy using implanted fiducial markers with daily portal imaging: assessment of prostate organ motion. *Int J Rad Bio Phys*. 2007;68:912–919.
25. Lin LIK. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45:255–268.
26. Lin LIK. Assay validation using the concordance correlation coefficient. *Biometrics*. 1992;48:599–604.
27. McBride GB. *A Proposal for Strength-of-Agreement Criteria for Lin's Concordance Correlation Coefficient*. NIWA Client Report: HAM2005-062; Hamilton, New Zealand: National Institute of Water & Atmospheric Research May 2005. Available online: <http://www.medcalc.org/download/pdf/McBride2005.pdf>, accessed on 16 March 2017.
28. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Stat Med*. 2012;31:3972–3981.
29. Martins WP, Nastri CO. Interpreting reproducibility results for ultrasound measurements. *Ultrasound Obstet Gynecol*. 2014;43:479–480.
30. Klein EE, Hanley J, Bayouth J, et al. Task group 142 report: quality assurance of medical accelerators. *Med Phys*. 2009;36:4197–4212.
31. Nichol AM, Brock KK, Lockwood GA, et al. A magnetic resonance imaging study of prostate deformation relative to implanted gold fiducial markers. *Int J Radiat Oncol Biol Phys*. 2007;67:48–56.
32. Klayton T, Price R, Buyyounouski MK, et al. Prostate bed motion during intensity modulated radiotherapy treatment. *Int J Radiat Oncol Biol Phys*. 2012;84:130–136.
33. deVan Water S, Valli L, Aluwini S, Lanconelli N, Heijmen B, Hoogeman M. Intrafraction prostate translations and rotations during hypofractionated robotic radiation surgery: dosimetric impact of correction strategies and margins. *Int J Radiat Oncol Biol Phys*. 2014; 88: 1154–1160.
34. International Organization for Standardization. *Guide to the Expression of Uncertainty in Measurement*. Geneva: International Organization for Standardization; 1995.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**Table S1.** List of randomly generated translational (mm) and rotational offsets (degrees).

**Table S2.** Combined uncertainty in reported translations (mm) and rotations (degrees) for Calypso and radiographic imaging. Uncertainties calculated according to the ISO GUM at the 95% confidence level.

## APPENDIX

### UNCERTAINTY IN MATCHING

The main contributor to the uncertainty in reported translations and rotations was rounding errors. The radiographic matching software only reports translations to the nearest millimeter, and although Calypso reports positions to one tenth of a millimeter, the Calypso results were rounded to the nearest millimeter prior to statistical analysis. Calypso only reports angles to the nearest whole degree. The radiographic matching software reports angles to tenths of a degree, but the values were rounded to the nearest degree prior to analysis.

For radiographic imaging, a second contributor to uncertainty is operator-dependent variation in matching. For kV orthogonal imaging, this was assessed by performing 20 independent matches on the same image set and calculating the standard deviation of reported translations and rotations. For translations, all 20 matches gave identical values (within the 1 mm resolution of the software), resulting in a standard deviation of 0 mm. For rotations, the standard deviations were  $0.4^\circ$ ,  $0.7^\circ$ , and  $0.4^\circ$  for pitch, roll, and yaw, respectively.

A similar set of 20 independent matches was performed on a single CBCT data set. As for the kV orthogonal images, there was a standard deviation of 0 mm for translations. For rotations, the standard deviations were less than  $0.1^\circ$  for all three angles.

Calypso rotations and translations are not subject to operator-dependent variation. However, there is random noise in the Calypso trace that contributes to uncertainty in the reported translations. This was assessed by assuming that the noise is normally distributed and estimating the standard deviation from the maximum excursion reported by the software. The estimated standard deviation was less than 0.1 mm in all three axes. Random noise is assumed not to contribute to uncertainty in reported rotations.

The combined uncertainties in reported translations and rotations were calculated according to the ISO GUM<sup>34</sup> and are summarized in Table S2.