Article:

Avila-Carrasco, C., Ruppel, M., Persad, R. et al. (2 more authors) (2020) Analytical versus data-driven approach of modelling brachytherapy needle deflection. IEEE Transactions on Medical Robotics and Bionics, 2 (4). pp. 519-528.

https://doi.org/10.1109/tmrb.2020.3032192

Analytical vs Data-driven Approach of Modelling Brachytherapy Needle Deflection

Journal:	*IEEE Transactions on Medical Robotics and Bionics*
Manuscript ID	Draft
Manuscript Type:	Original Articles
Date Submitted by the Author:	n/a
Complete List of Authors:	AVILA CARRASCO, CAROLINA; Bristol Robotics Laboratory, RUPPEL, MIRJANA; Bristol Robotics Laboratory PERSAD, RAJENDRA; North Bristol NHS Trust BAHL, AMIT; University Hospitals Bristol NHS Foundation Trust DOGRAMADZI, SANJA; The University of Sheffield, Department of Automatic Control and Systems Engineering
Keywords:	Brachytherapy, Imageless needle tip tracking, Needle deflection model, Multilayer perceptron, Artificial Neural Network
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
Video_Plastic needle_Ver Plane.mp4 Video_Plastic needle_Hor plane.mp4 Video_Ti needle_Ver Plane.mp4 Video_Ti needle_Hor plane.mp4	

Analytical vs Data-driven Approach of Modelling Brachytherapy Needle Deflection

Carolina Avila-Carrasco, Mirjana Ruppel, Rajendra Persad, Amit Bahl, Sanja Dogramadzi

Abstract— **This research is motivated by the need of real-time needle tracking solutions in brachytherapy procedures for improving targeting accuracy. We compare two different modelling approaches to estimate brachytherapy 3D needle deflection during insertion into soft tissue from reaction forces and moments measured at the base of the needle: an analytical model based on beam deflection theory and, a data-driven model using a multilayer perceptron artificial neural network (ANN). Verification of the analytical model as well as training, validation, and testing of the ANN model were performed with experimental data obtained from over 120 insertion tests into gelatine tissue phantoms including a variety of needle types and tissue properties. The ANN model has lower prediction errors and is more robust to changes in testing conditions, with accurate predictions in 3 out of 4 tested scenarios; whereas the analytical model predictions are not statistically comparable to ground truth values in any of the tested scenarios. ANN models show a big potential for online 3D tracking of brachytherapy needles in a clinical context in comparison with beam theory analytical models. A simple neural network trained with numerous needle insertions into representative biological soft tissue could estimate needle tip position with submillimetre accuracy.**

Index Terms— **Brachytherapy, Imageless needle tip tracking, Needle deflection model, Multilayer perceptron, Artificial Neural Network**

I. Introduction

Brachytherapy is a localized radiotherapy procedure to treat prostate cancer through targeted radiation applied to the affected tissue using special needles (Fig. 1). Among existing brachytherapy techniques, low-dose rate (LDR) brachytherapy consists in permanent placement of multiple radioactive seeds into the prostate and surrounding tissue [4]. The implantation locations of the seeds and their radiation dose are calculated through computer-assisted preoperative planning with multiplane clinical images of the prostate anatomical region. After the plan is completed, radioactive sources are loaded on the needles and these are inserted into the prostate gland up to the planned depths using a guiding external grid template. Needle insertion procedure is done manually using transrectal ultrasound (TRUS) image for guidance. This visual feedback helps the surgeon to track the needle trajectory and tip location during the insertion, but does not provide a quantitative measurement of the needle tip location nor of the deviation from the planned trajectory.

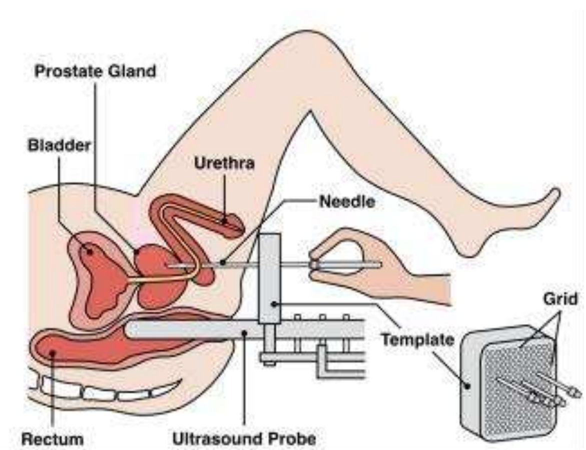


Fig. 1. Brachytherapy needle insertion procedure for prostate cancer treatment. In LDR brachytherapy radioactive seeds are permanently implanted in the prostate gland using needles inserted through a guiding grid template. The procedure is image-guided by transrectal ultrasound (TRUS).

Targeting accuracy is a critical factor for the clinical effectiveness of brachytherapy procedures since a misplacement of the implanted radiation sources from the planned optimal locations would alter the radiation dose received by the tumour [5, 20]. Needle deflection and soft tissue deformation during the needle insertion process affect targeting accuracy producing an average deviation of the implanted seeds of 3-6mm from the intended target [5, 11, 15, and 21]. According to a recent study, positioning error thresholds to prevent a significant change of the radiation dose in prostate treatment range from 2 to 5 mm [20], confirming the importance of targeting accuracy.

Real-time tracking of the needle tip trajectory during the insertion procedure and quantification of the deviation from the planned path can help to minimize seed placement errors. By

This work was partially funded by Above & Beyond, Registered Charity No. 11709.
C. Avila-Carrasco is with the Bristol Robotics Laboratory – UWE, Bristol, UK (e-mail: carolina.avila@yahoo.es).
M. Ruppel is with the Bristol Robotics Laboratory – UWE, Bristol, UK (e-mail: mirjana.ruppel@uwe.ac.uk).
R. Persad is with the North Bristol NHS Trust, Bristol, UK (e-mail: raj.persad@nbt.nhs.uk).
A. Bahl is with the University Hospitals Bristol NHS Trust, Bristol, UK (e-mail: amitbahl@doctors.org.uk).
S. Dogramadzi is with the Department of Automatic Control and Systems Engineering, The University of Sheffield (e-mail: s.dogramadzi@sheffield.ac.uk).

knowing the deviation from the planned path at all times, the surgeon would be able to apply the necessary corrections to reach the target position or adjust the treatment plan accordingly.

For the development of real-time needle tracking solutions for brachytherapy procedures, two different approaches have been considered by the research community: image-guided tracking using computer vision technology and imageless tracking using non-vision based sensors. Image-guided tracking is a complex and computationally expensive process that involves image acquisition in more than one plane, image processing and robust computer vision algorithms in order to identify the needle tip in the intraoperative TRUS images and calculate its 3D position. Real-time 3D tracking of needle insertion using the low resolution ultrasound (US) images is still a challenge, mainly due to the difficulty in the needle image segmentation. Continuous visualization of the needle tip on different US image planes during the whole insertion procedure is another technical challenge for the design of these image-guided systems.

On the other hand, imageless tracking technologies are seen as an alternative solution to computer vision systems for real-time needle tracking due to their lower computational demands and their capability for high frequency sensor data acquisition. Among imageless technologies, needle tip location could be directly measured using electromagnetic sensor coils placed inside the needle [1, 23]. These sensors allow real-time 3D tracking; however, their measurement accuracy can be significantly affected by their limited sensing field, with accuracy dropping over a certain distance to the magnetic field generator, as well as their sensitivity to the presence of metals, which can produce significant distortion and drift in the signal. Moreover, the use of sensors inserted within the needle could alter the deflection behaviour of the needle, introducing a new source of error to the needle trajectory. An alternative, less invasive, approach is to indirectly measure needle deflection using predictive needle-tissue interaction models with real-time input data obtained through force sensors placed at the base of the needle. This last approach is the one selected in the present study; an indirect measurement of needle tip deflection by using real-time force data measured by a load cell during needle insertion combined with a predictive model that calculates needle deflection using force data as inputs.

Most published needle-tissue interaction models are analytical and calculate needle deflection using different mechanical principles: Webster *et al.* [25] proposed a nonholonomic kinematic model where needle motion is compared to that of a bicycle with a fixed curved trajectory. The application of this type of model requires a previous experimental characterization of each needle and tissue combination in order to fit the model parameters. A different approach was proposed by Glozman *et al.* [8], who modelled the needle as a series of linear beam elements supported by virtual linear springs simulating needle-tissue interaction forces. The stiffness coefficients of the springs are obtained experimentally through image processing of the needle shape and calculation of the displacement of the virtual spring points from the reference straight line. This method is suitable for highly flexible needles which can change shape along the path. It requires clinical images to adjust the model parameters from the needle shape, adding more complexity to the estimation of needle deflection. Goksel *et al.* [9] developed an 'angular springs' model where the needle is split into small rigid rods connected by spring-loaded joints. Consecutive needle segments bend and twist relative to each other when exposed to external loads. Two rotational springs at each joint simulate the internal reaction torques or resistance of the needle shaft to bend or twist. The model spring constants are obtained by fitting of experimental data. Energy-based formulations have also been used to model needle deflection behaviour: Misra et al. [18] created a 2D model of flexible bevel tip needles where the total energy of the system is expressed in terms of the transverse and axial deflections of the needle. Inputs to the model are the material and geometric properties of both the needle and the tissue, therefore a previous characterization of the tissue properties is required.

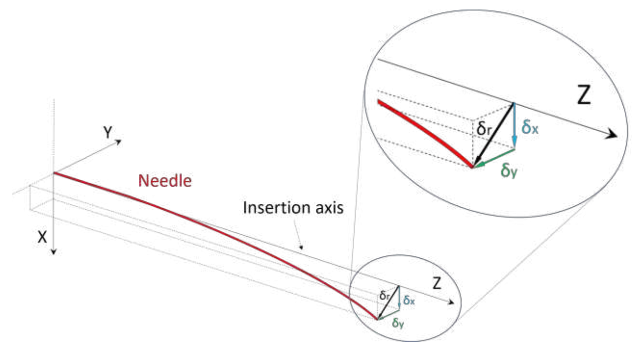


Fig. 2. Graphical representation of needle deflection. Total needle deflection in our study δr is defined as the Euclidean distance of the needle tip from the needle insertion axis; it is the resultant of perpendicular deflections δx (vertical) and δy (horizontal).

The above analytical models were mainly defined for highly flexible needles which experience significant deflection and curvature. However, brachytherapy needles are only moderately flexible, experiencing small deflections in respect to their length. This is why mechanical models based on Euler-Bernoulli beam theory have shown a reasonably good performance for estimating deflection of brachytherapy needles [1, 12-15]. Beam-theory models establish an equivalence between the needle and a cantilever beam and define needle deflection as a function of the loads acting on the needle during insertion into soft tissue. A key assumption for the applicability of beam-theory equations is that beam deflections must be significantly low with regard to the total beam length. Needles used for brachytherapy procedures are rigid enough and experience sufficiently small deflections to meet beam-theory requirements. In fact, beam-theory models have shown deflection estimation errors below 1mm when applied to *in-vitro* needle insertion experiments, using both phantoms and biological tissue [1, 13]. A practical advantage of beam-theory models against the analytical models cited above is that they do not require previous experimental characterization of tissue mechanical properties, thus allowing a more straightforward application. Needle deflection is calculated by applying Euler-

Bernoulli equations to the proposed needle-tissue interaction model with the only input variables being the reaction forces and moments at the base of the needle and the insertion depth. Reaction loads can be measured in real time during needle insertion using a load cell attached to the needle base support.

Nevertheless, despite the reasonable performance shown by analytical models in some studies, the heterogeneous and multilayer nature of biological tissue limits the accuracy and clinical applicability of this type of models which are built on assumptions that simplify the behaviour of the needle and surrounding tissue. Data-driven models represent an alternative approach to estimate needle deflection under highly nonlinear conditions typical for heterogeneous physiological tissues [22]. However, the accuracy of these models depends on the size and variety of the available database of observations. In this paper, we compare the performance of two modelling approaches for estimating brachytherapy needle deflection during insertion into soft tissue phantoms without using image guidance: an analytical model based on beam deflection theory and a data-driven model using a multilayer perceptron artificial neural network. In both cases, input parameters are the reaction forces and moments measured at the base of the needle, the insertion depth, and needle structural properties. Outputs from both models are needle tip deflections in two orthogonal axes (Fig. 2), allowing for 3D tracking of the tip position during insertion. Accuracy of both models was assessed using ground truth measurements of needle tip deflections obtained from continuous optical tracking of the needle insertion in two perpendicular planes.

II. Methodology

A. Beam-Theory model

The model estimates needle tip deflection in two orthogonal axes by applying static beam deflection equations to two perpendicular bending planes (vertical or XZ and, horizontal or YZ). Total needle tip deflection δr has been defined as the Euclidean distance of the needle tip from the needle insertion axis in the perpendicular plane to the insertion axis. Total deflection δr can be split into perpendicular components δx and δy, which quantify needle tip deflections along vertical (x) and horizontal (y) axes, respectively (Fig. 2).

Fig. 3 shows a graphical representation of the model in both deflection planes considered for analysis. L represents the total length of the needle, d is the inserted length and a is the length outside the tissue (L-d). Reaction loads at the needle's base are force Frx and torque Try in the vertical plane XZ and force Fry and torque Trx in the horizontal plane YZ (only those reaction loads contributing to needle bending are considered in beam-deflection models). Needle-tissue interaction load is modelled with a uniform force distribution along the inserted portion of the needle having a force intensity per unit length of qx in plane XY and qy in plane YZ. Additionally, a single-point force Ftx acting on the tip of the needle along the vertical X axis accounts for the needle-tissue interaction resulting from the tissue cutting force. This point force was only considered in the vertical plane for consistency with the needle behaviour observed in our experimental tests, with predominantly larger vertical deflection than horizontal deflection. Proposed needle-tissue interaction load profiles are based on previously published beam-theory models for brachytherapy needles [1, 12-15]. In a previous research work [2] we compared the performance of the analytical model with different interaction load profiles using the same experimental data and the selected profiles shown in Fig. 3 are those producing the lowest mean absolute prediction error on each deflection plane.

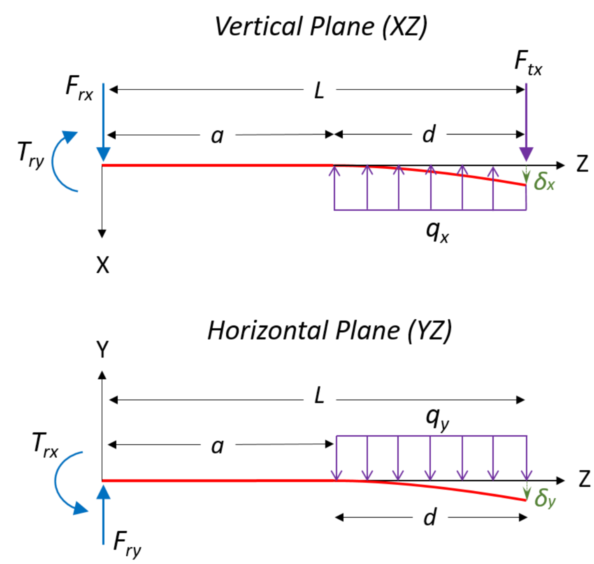


Fig. 3. Load profiles characterising the proposed beam-theory model. Top: Forces acting on the vertical plane (XZ); Bottom: Forces acting on the horizontal plane (YZ).

Model inputs are: 1) needle intrinsic parameters: Young's Modulus (E), area moment of inertia (I), and total length of the needle (L), 2) needle insertion depth (d), and 3) reaction forces and torques measured at the base of the needle that contribute to needle deflection (Frx, Try, Fr and, Trx). Model outputs are the vertical, δx, and horizontal, δy, needle tip deflections.

The basic differential equation of the deflection curve of a beam which relates bending moment M and deflection $v(z)$ at a distance z from the base can be written as (1), where E is the Young's modulus and I is the area moment of inertia of the beam.

$$M = EI\frac{d^2v}{dz^2} \quad (1)$$

For each deflection plane, needle tip deflection formula is obtained from double integration of the bending moment's equation; using the superposition principle when more than one type of load is acting on the needle [7].

For a cantilever beam subject to a uniform load distribution of intensity q along a distal portion d of the total length L, the formula for the tip deflection in a generic bending plane is:

$$v_q(z = L) = \delta_q = \frac{q}{24EI}(3L^4 - 4a^3L + a^4) \quad (2)$$

Likewise, the formula for tip deflection for a cantilever beam under a point force F_t at the tip is:

$$v_{Ft}(z = L) = \delta_{Ft} = \frac{F_t L^3}{3EI} \quad (3)$$

Needle tip deflection δx in the vertical plane can be obtained by applying the superposition method. Total vertical deflection is calculated with equation (4) as the sum of the vertical deflections due to the uniform load distribution qx (2) and the point force Ftx (3). The force intensity value qx and point load at the tip Ftx are previously obtained from the force and moment equilibrium equations at the needle, where the rest of parameters are known (reaction loads at the base of the needle are given measurements).

$$\delta_x = \delta_{qx} + \delta_{Ftx} = \frac{q_x}{24EI}(3L^4 - 4a^3L + a^4) + \frac{F_{tx}L^3}{3EI} \quad (4)$$

Needle tip deflection δy in the horizontal plane can be calculated with equation (5), where the force intensity value qy is first obtained from the force and moment equilibrium equations at the needle.

$$\delta_y = \frac{q_y}{24EI}(3L^4 - 4a^3L + a^4) \quad (5)$$

B. Neural-Network model

Multilayer perceptron (MLP) artificial neural networks are known as powerful function approximators, able to fit any input-output mapping problem. A MLP network with one hidden layer of neurons using continuous nonlinear sigmoid activation functions and one output layer with linear activation function neurons can approximate any continuous function to arbitrary precision, provided the network has a sufficiently large number of hidden neurons [3]. Another advantage of MLP networks is that they are good at generalization even for small datasets providing reasonably accurate predictions for datasets independent of those used for training and validation. For these reasons, a multilayer perceptron feed-forward artificial neural network (ANN) was selected as a suitable data learning approach to estimate needle deflection from a set of known parameters characterizing needle insertion behaviour.

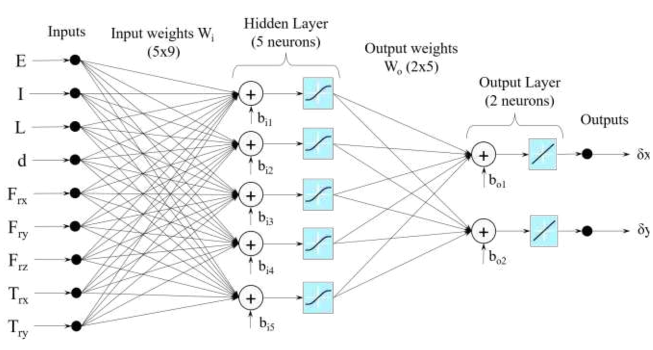


Fig. 4. Architecture of the proposed multilayer perceptron feed-forward artificial neural network for the data-driven model. The network has one hidden layer with 5 neurons using sigmoid functions and one output layer with 2 linear neurons.

The proposed ANN has 9 inputs, 2 outputs and one hidden layer with 5 neurons. Neurons in the hidden layer use hyperbolic tangent sigmoid activation functions while those in the output layer use linear activation functions (Fig. 4). Inputs are: needle Young's Modulus (E), area moment of inertia (I), total length of the needle (L), needle insertion depth (d), reaction forces measured at the base of the needle (Frx, Fry, Frz), and reaction torques contributing to needle deflection (Trx, Try). Outputs are the vertical, δx, and horizontal, δy, needle tip deflections. The network was trained with data obtained from multiple needle insertion tests into soft tissue phantoms. Training was performed using the 'Levenberg-Marquardt' backpropagation algorithm. A ten-fold cross-validation was applied to the training process which consists of randomly splitting the database into 10 data subsets of equal size and train the network 10 times using a different subset as the validation data for each training run. From the 10 generated network models, the one with the best validation performance was selected for further evaluation with test data. Network training and validation performances were quantified by the mean squared error between the network outputs and the target values which are the ground truth deflections. Matlab Neural Network Toolbox [10] was used to generate and train the ANNs.

TABLE I
DATA DISTRIBUTION FOR THE 4 ANN MODEL VARIANTS GENERATED IN THE STUDY (CHARACTERISTICS AND SAMPLE SIZE OF EACH DATA SUBSET)

ANN Model	Training Data		Validation Data		Test Data	
	Description	Data points	Description	Data points	Description	Data points
Trial 1	10% and 20% Tissue Density	369	10% and 20% Tissue Density	41	5% Tissue Density	80
Trial 2	5% and 10% Tissue Density + 20% Density with no skin	369	5% and 10% Tissue Density + 20% Density with no skin	41	20% Tissue Density + Skin	80
Trial 3	Insertion depths ≤ 75mm	324	Insertion depths ≤ 75mm	36	Insertion depths = 97mm	123
Trial 4	Random	394	Random	49	Random	49

A database of 492 points was created from 123 needle insertion tests performed under a variety of conditions (described in section II.C). In order to evaluate the generalization capability of the proposed ANN, four different ANN model variants were obtained by changing the distribution of training, validation and test data subsets from the original database (Table I). For the first model, selected test data were all points from the lowest tissue hardness (density 5%) while training and validation data were points obtained from harder tissue samples (densities 10% and 20%). For the second model trial, selected test data were all points from tissue samples with density 20% and additional skin layer, while training and validation data were those generated from the rest of tested tissue properties. These two first scenarios would allow assessing the ANN's performance for different tissue properties to those used for training the network. In the third model version, test data were all points for the maximum insertion depth from each insertion test whereas training and validation data belonged to shorter insertion depths. This would allow investigating the applicability of the ANN to insertion depths longer than the ones used for training. In the fourth trial, test data were 10% of randomly selected points from the database. For each tested ANN variant, the test data subset was first removed from the original dataset and a ten-fold cross-validation training was applied to the remaining database. Performance of the resultant ANN models was assessed on the corresponding test data subsets. Performance of both analytical and ANN modelling approaches was compared through *two-sample t-tests* between ground truth deflection and each model predicted deflections as well as through comparison of their respective mean absolute prediction errors (MAE) for the four different test data subsets.

C. *Data acquisition needle insertion experiments*

A total of 123 needle insertion tests into soft tissue phantoms were carried out with the aim to gather data for training the ANN as well as for evaluation of the performance of both, the analytical and the ANN models.

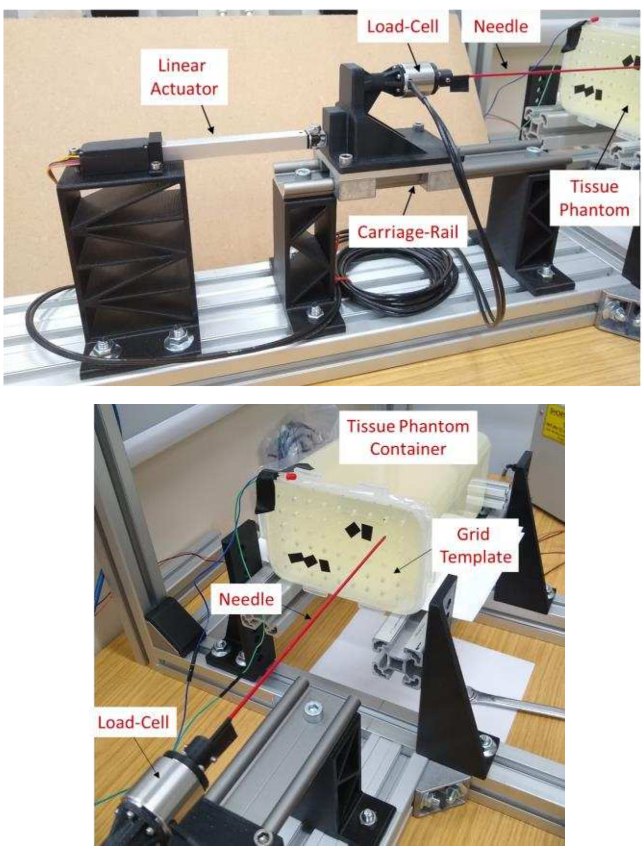


Fig 5. Experimental set-up for needle insertion tests. Top image: Mechatronic device for needle insertion and retraction using a linear actuator attached to a carriage holding the needle. Bottom image: Detail of the needle mounted on the insertion device and tissue phantom container. A load-cell is fixed to the base of the needle holder and measures reaction forces and torques at the base of the needle.

A custom mechatronic system was built to perform needle insertion and retraction using a linear actuator with a maximum stroke of 98 mm (Firgelli L16-100-63-12P). Displacement of the needle is controlled by an Arduino Uno Rev.3 board combined with an Arduino Motor Shield to control speed and direction of the linear actuator. Insertion depth is continuously monitored through the linear actuator built-in potentiometer while insertion speed can be modified using PWM. A carriage-rail system allows the displacement of the needle ensuring a precise alignment to the insertion axis (Fig. 5).

A 6-axis load cell (Interface 6A27) is attached to the base of the needle holder and used for real-time measurement of reaction forces during the insertion tests. It has a load capacity of 200N for traction and compression forces, 50N for shear forces, and 1Nm for all torques. It is connected to a dedicated data acquisition device including signal conditioning and amplifier, with synchronized sampling at 24 bit resolution (Interface BX8-HD44 BlueDAQ series). The BlueDAQ data logging software reads and saves the load cell data referenced to the base of the needle, as well as the position of the linear actuator through a shared connection with the Arduino Uno board. This way, a synchronised reading between reaction loads at the base of the needle and the insertion depth is obtained. Data acquisition frequency was set to 100 Hz for the experiments in this study, which is considered enough for continuous tracking of the needle at the tested insertion speeds.

Ground truth needle deflections along insertion tests were measured using optical tracking with video cameras. This method has been previously used in research studies to measure needle deflection when tissue samples are transparent [14-15, 24-25]. The choice of optical tracking combined with transparent tissue phantoms is motivated by the simplicity of the experimental setup and the straightforward identification of the needle profile and tip on the standard images compared to US images. Some researchers have alternatively used electromagnetic sensors placed inside the needle [1, 23] but as mentioned in the paper introduction, the accuracy of this type of sensors can be affected by their limited sensing field and electromagnetic interferences in the surrounding environment.

Needle insertion was optically tracked with two video cameras (PointGrey –Grasshopper3 GS3-U3-41CEC) fixed to a frame structure and with their axes perpendicular to the vertical (XZ) and horizontal (YZ) needle insertion planes respectively (Fig. 6). This allows measuring needle deflection in two orthogonal planes which are also coincident with the load-cell coordinate system XZ and YZ planes, enabling direct comparison between ground truth measurements and model estimations. Position of the cameras relative to the needle insertion mechatronic device was fixed, therefore the distance between each camera and the associated needle insertion plane was kept constant in all experiments. To perform different insertions within the same tissue sample we moved the tissue phantom up/down or left/right with regard to the needle insertion device, resulting in a variable tissue thickness between the cameras and the needle insertion plane. Video acquisition of both cameras was synchronised at 20 frames per second with an image resolution of 1024x1024 pixels using the Matlab

Image Acquisition Toolbox [16]. In order to maximize the image resolution, each camera field of view was adjusted to the region of interest, covering the needle length before insertion plus the insertion depth. Before deflection measurements, image calibration was performed for each camera to find the ratio between the number of pixels and distance in mm corresponding to the corresponding needle insertion plane. For each deflection plane, an image calibration was performed by placing a ruler coincident with the needle plane along the needle axis. We recorded an image with the camera using the same setting parameters as for the insertion tests. Finally, the ratio mm/pixel was obtained by measuring the equivalent distance in pixels between 2 points in the ruler at a known distance in mm. This image calibration was done using Matlab image processing tools and resulted in an average ground truth measuring resolution of 0.33mm/pixel for XZ plane (vertical deflection δx) and 0.40mm/pixel for YZ plane (horizontal deflection δy). Repeatability of the ground truth method was assessed by repeating needle deflection measurements in 20 tests. Average absolute differences between repeated measurements were 0.21 mm for images of plane XZ and 0.25mm for images of plane YZ, both lower than the established resolution of each camera.

Video recording was not synchronized with load-cell and insertion depth data but they were paired later by using timestamps on the video files and data files and identifying the start of the insertion test with an LED flash captured by both cameras.

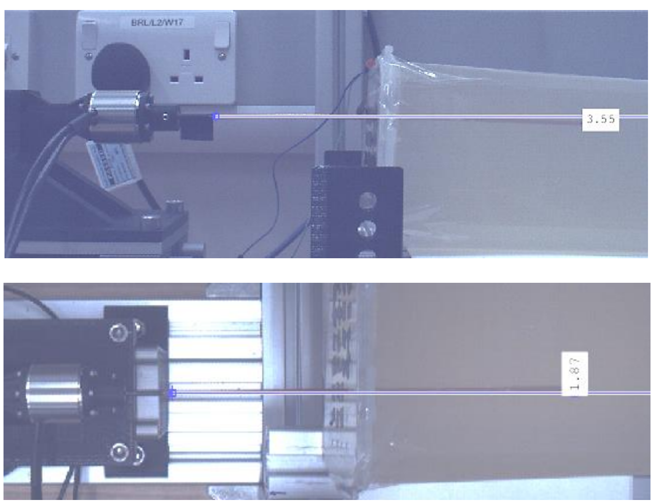


Fig. 6. Example of ground truth deflection measurement. Tip deflection values in vertical and horizontal planes for a certain insertion depth are obtained from image processing of the corresponding video-frames. Top image corresponds to the camera for the vertical plane (XZ) and bottom image corresponds to the horizontal plane camera (YZ).

Ground truth needle tip deflection was measured at four specific insertion depths for each experiment: 25mm, 50mm, 75mm and 97mm. For each insertion depth, the corresponding frames of each deflection plane were extracted from the recorded video-files and processed by a Matlab script to measure vertical (δx) and horizontal (δy) deflection values. The script uses image processing tools [17] to measure needle deflection from two input points on each image: the base of the needle and the needle tip. These two points need to be manually marked on an image display window and the program calculates the deflection value as the perpendicular distance between the needle tip and needle insertion axis (Fig. 6). Ground truth deflection values for each insertion depth are measured as net deflections, relative to the initial position of the needle tip before the start of insertion.

TABLE II
SUMMARY OF NEEDLE INSERTION TESTS PERFORMED IN THE STUDY GROUPED PER COMBINATIONS OF NEEDLE TYPE, TISSUE DENSITY AND USE OF ADDITIONAL SKIN TISSUE LAYER. FOR EACH GROUP, TESTS WERE EQUALLY SPLIT INTO TWO AVERAGE INSERTION SPEEDS OF 10MM/S AND 15MM/S

Needle Type	Tissue Density	Skin Layer	Number of Tests
Plastic	5%	No	10
	10%	No	9
		Yes	18
	20%	No	10
		Yes	10
Titanium	5%	No	10
	10%	No	20
		Yes	16
	20%	No	10
		Yes	10

In order to have a varied sample database that allows testing the generalization performance of both needle deflection models, different combinations of needle type, tissue properties and insertion speeds were considered in the experiments design (see Table II). Tissue phantoms were made of porcine gelatine at specific concentrations. Three different density grades were tested by mixing gelatine and water in concentrations of 5%, 10% and 20% with phantom tissue hardness increasing with gelatine concentration. Tissue samples were built and tested inside a transparent plastic container and a plastic plate with an array of holes was placed over the insertion face, simulating a brachytherapy guiding grid template (see Fig. 5). This ensures the needle is always inserted into a different tissue region. To emulate the skin layer, we performed some experiments with an addition of a 5mm thick layer of silicone (Superflab Plastic Bolus Material) on top of the gelatine tissue.

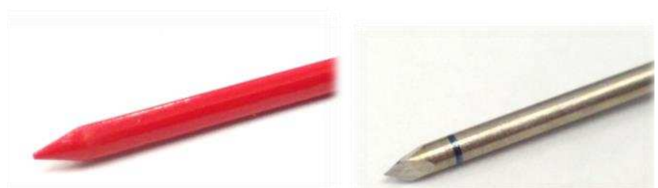


Fig. 7. Detail of the tip geometry of the brachytherapy needles used in the study. Left: Conical tip plastic needle with 2mm diameter. Right: Sharp trocar tip titanium needle with 1.65mm diameter.

Two types of symmetric tip brachytherapy needles (Varian Medical Systems, Fig. 7) were used in the study: 1) Plastic Needles (PEEK material) of 2 mm diameter and 200 mm length with conical tip geometry. 2) Titanium needles of 16 G (1.65 mm outer diameter) with two lengths of 200 mm and 250 mm and a sharp trocar tip geometry. Young's Moduli of tested needles were previously obtained experimentally from specific needle deflection tests [2] with average values of 27,65 MPa for the plastic needles and 107,74 MPa for the titanium needles. Tests were performed at two different average linear speed values of 10 mm/s and 15mm/s for each needle-tissue type combination.

For each insertion test, measured experimental data were model input parameters like insertion depth and reaction loads at the base of the needle as well as ground truth deflection values at the specified insertion depths. These data were used for evaluation of the analytical model performance and in the case of the ANN model, to train the network first and test its performance later.

III. Results and Discussion

Performance of both analytical and ANN models was compared for the test data subsets described in Table I. Results for total absolute needle tip deflection (|δr|) are shown in Table III, which includes: 1) average and standard deviation values of ground truth measurements and model predictions for each data subset; 2) MAEs given by both models; 3) results from the two-sample t-tests between measured and predicted deflections.

TABLE III
COMPARISON OF ANALYTICAL AND ANN MODELS PERFORMANCE FOR PREDICTION OF TOTAL ABSOLUTE TIP DEFLECTIONS IN FOUR DIFFERENT TEST DATA SUBSETS

Results for Total Absolute Needle Tip Deflection (|δr|)					
Test Data Subset	Average Deflection ± STD (mm)			MAE ± STD (mm)	
	Ground Truth	Analytical Model [a]	ANN Model [a]	Analytical Model	ANN Model
Trial 1	0.78 ± 0.61	1.14 ± 0.55 *	1.37 ± 0.57 *	0.50 ± 0.29	0.62 ± 0.41
Trial 2	1.15 ± 0.93	7.49 ± 4.15 *	1.19 ± 0.88	6.35 ± 3.47	0.32 ± 0.28
Trial 3	2.00 ± 1.46	6.13 ± 4.65 *	1.84 ± 1.03	4.23 ± 3.82	0.50 ± 0.82
Trial 4	1.13 ± 0.84	4.13 ± 3.51 *	0.99 ± 0.81	3.03 ± 3.04	0.29 ± 0.22

[a] * Hypothesis of equality of means between ground truth and predicted deflections is rejected (*Two-sample t-tests*)

Fig. 8 shows a box-plot comparison of ground truth total absolute deflections with analytical and ANN model predictions for the same test data subsets. Results have been grouped by the type of needle and the tissue hardness (proportional to density value), which are those factors with a significant effect on the needle deflection behaviour according to a previous research work [2].

The ANN model has lower prediction errors than the analytical model in all test data subsets except for Trial 1, where both models have a comparable performance. Results from the two-sample t-tests reveal that analytical model predictions are significantly different to ground truth deflections in all test datasets, whilst ANN model deflection estimations are statistically equivalent to ground truth values in all cases except for Trial 1.

The ANN modelling approach shows a consistent performance across all tested scenarios with a MAE below 1 mm. Excluding Trial 4, where test data were randomly selected from the overall database, in the rest of trials, data used to test the network corresponded to different tissue properties (Trials 1 and 2) or different insertion depths (Trial 3) from the data points used to train the network. The low prediction errors obtained in all trials suggest a good generalization performance of the proposed ANN approach.

When looking at how different factors like needle type or tissue density may affect both models' performance, we observe a more robust behaviour of the ANN approach compared to the beam-type analytical model (see Trials 3 and 4 box plots in Fig. 8). Whereas ANN predicted deflections have very close distributions to ground truth values regardless of the type of needle and soft tissue properties, a much less consistent behaviour is observed in the analytical model's performance across different groups. The beam-type model produces generally larger prediction errors for plastic needles than for metal needles, and its predictions get worse with increasing tissue hardness with both types of needles. The analytical model only has mean prediction errors below 1 mm with test data belonging to the lowest tissue density (Trial 1). However, needle deflection values for the softest tissue samples are of the same order of magnitude as the model prediction errors, both for the analytical and ANN approaches. This makes impractical any assessment of the models accuracy for those test conditions where average actual needle deflections are in the submillimetre range.

The reason why needle deflections measured in this study are generally small is because insertion experiments were performed with symmetric tip needles. These needles have minimal deflections compared to bevel tip steerable needles, which can bend over 10 mm for insertion depths of 100 mm [19, 24]. Future work will include a verification of the two proposed modelling approaches for bevel tip needles. Despite this limitation, results from this first comparative study are able to discriminate between both modelling approaches and indicate a better and more robust performance by the ANN model.

Apart from needle intrinsic parameters characterizing the mechanical behaviour of the needle, input variables in both models are the measured reaction loads at the base of the needle. Although 2 different insertion speeds were tested in the study, this was not considered as input variable. The analytical model is based in beam deflection theory which considers a static equilibrium of all forces and moments acting on the needle, therefore it does not include dynamic parameters. With regards to the data-driven ANN model, insertion speed was not considered as an input variable because we didn't find a significant influence of this parameter on needle deflection on a previous statistical analysis of the data using Analysis of Variance (ANOVA) and non-parametric Kruskal-Wallis tests [2]. There is some discrepancy among different published studies with regard to the effect of needle insertion speed on deflection. Some authors did not find a significant effect [15, 25], while others reported an opposite outcome with a significant influence of speed in needle deflection behaviour [24]. Discrepancy among different studies is probably due to the different range of speed values used across studies as well as different types of needles. It could be that the difference between average speeds selected for our study (10mm/s and 15mm/s) is not large enough to produce significant changes on the needle deflection. However, it is worth noting that we included the compression reaction force (*Frz*) as one of the input parameters of our ANN model and this parameter has a strong correlation with the insertion speed. Due to the viscoelastic nature of the soft tissue phantoms, faster needle insertions on the same tissue produce an increase of the compression force on the needle. We included this compression force parameter as one of the inputs of the ANN model because we observed that its inclusion resulted in a better prediction performance by the network.

Results from this study reflect a limited capacity of beam-theory analytical models for prediction of needle deflection during insertion into soft tissue. The analytical model proposed in this study provides reasonable estimations only for the softest monolayer tissue samples (5% density with no skin layer). Harder tissue phantoms and the addition of skin layer yield more complex needle-tissue interactions which are not properly captured by the model. In the same way as other types of analytical models, beam-theory models are built on assumptions that simplify the behaviour of the needle and surrounding tissue. Beam-theory models simplify the definition of needle-tissue interaction by defining an arbitrary force distribution profile. Identification of which load profile would be the best fit for each insertion point in time is a technical challenge, since direct measurement of needle-tissue interaction forces is not possible. Beam-theory models use specific shear force distribution profiles to define needle-tissue interaction forces responsible of needle bending. These load distribution profiles are generally arbitrary (e.g. uniform force distribution, triangular force distribution, point loads, etc.) and are normally kept fixed along the insertion process, with only the magnitude and sign of the forces changing in real time to meet force and moment equilibrium equations. Although this simplification of the interaction forces might work well for certain needle-tissue combinations, it is quite unlikely that a simplified force distribution profile will be representative of all possible needle-tissue interactions. Also, even for the same type of needle and tissue combination, the shape of the force profile will most likely change throughout the insertion due to the dynamic nature of the process and the heterogeneous nature of biological tissue. The challenge of defining representative forces profiles for any particular needle-tissue interaction at any time during the needle insertion process results in beam-theory needle deflection models being unreliable for clinical application.

In contrast to analytical models, data-driven models like the MLP neural network defined in our study have the potential to capture the complexity of the heterogeneous and anisotropic nature of human soft tissues and organs as long as they are trained with a sufficiently large data base of experimental data from representative biological tissue. More complex ANN architectures could be proposed to predict needle deflection from a larger set of inputs and a larger number of hidden layers and neurons; however, we do not think that increasing the complexity of the neural network would significantly improve the accuracy of the model predictions in this case. A simple multilayer perceptron (MLP) architecture like the one proposed in the study is a typical ANN architecture used in non-linear

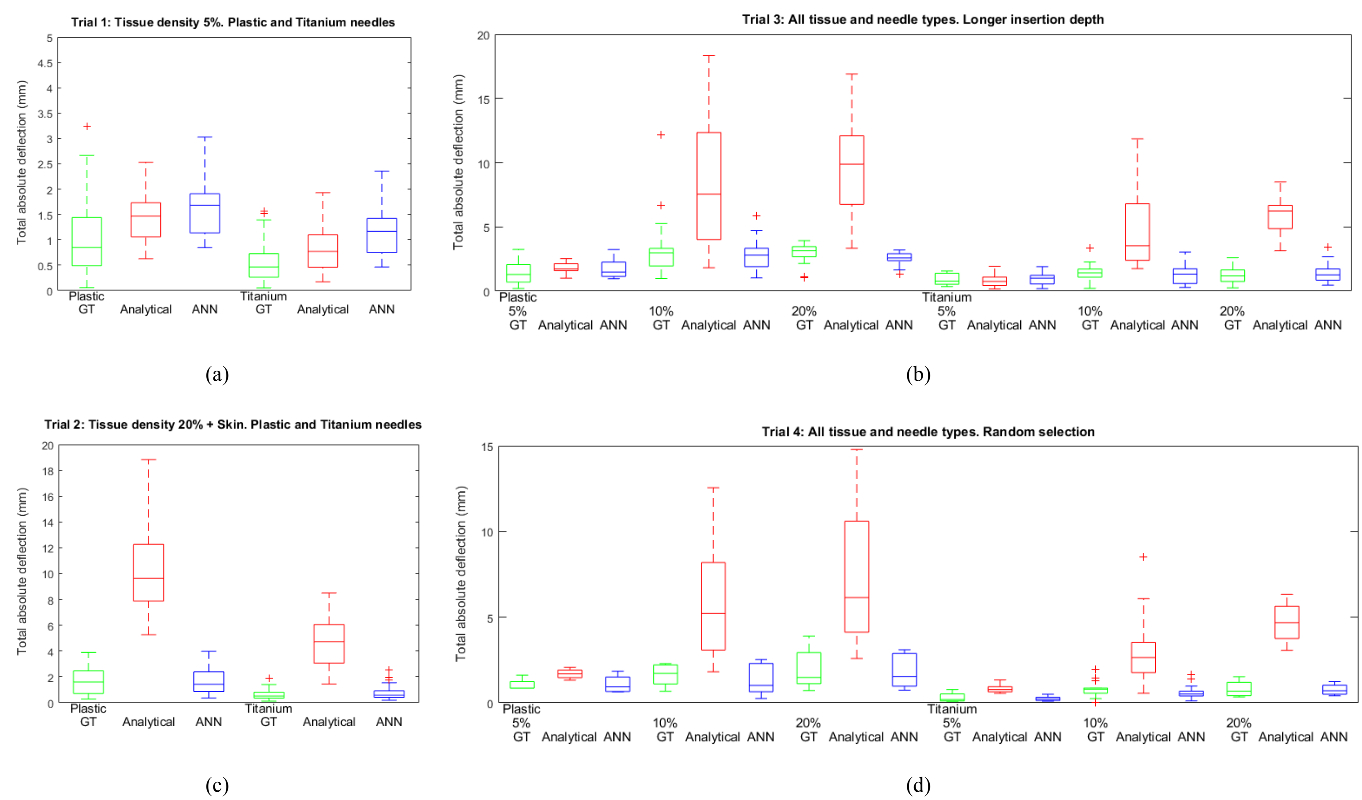


Fig 8. Box-plot distributions of ground truth (GT) and predicted (Analytical & ANN) total absolute needle tip deflection (|δr|) grouped by needle type (Plastic / Titanium) and soft tissue hardness (5, 10 and 20% densities). Individual plots show the absolute deflections for each of the test data subsets defined in Table I (Trial 1 (a), Trial 2 (b), Trial 3 (c) and Trial 4 (d)). Box-plots include data between the first and third quartiles, whiskers extend up to ±2.7σ and red crosses represent outliers. Box colours are used for better clarity to distinguish GT (Green), Analytical (Red) and ANN (Blue) deflections.

regression applications. Feedforward MLP neural networks are well-known powerful function approximators because of their ability to fit any finite input-output mapping problem. A two-layer MLP like the one used in the study, with one hidden layer of neurons using continuous nonlinear sigmoid transfer functions and one output layer with linear function neurons, can approximate any continuous function to arbitrary precision, provided the network has a sufficiently large number of hidden neurons [3]. The number of neurons in the hidden layer is a key aspect in the design of MLP since a low number may not be enough to learn properly the underlying function whereas a large number increases the risk of overfitting the data. In the design of our MLP model sizes between five and ten neurons for the hidden layer were explored and it was observed that the accuracy of network predictions did not significantly increase when using more than five neurons. Therefore, five neurons is the minimum acceptable size for the hidden layer in our MLP model to provide fairly accurate predictions while avoiding overfitting, thus achieving a good generalization performance.

Applicability of any ANN model relies on an adequate training which improves with the size and diversity of the dataset used for learning. In this study we created a starting database of 492 points, from which over 300 points were used for training the neural network model on each trial. Although more data points could have been extracted from a larger number of insertion depth points, we think that 4 data points per needle insertion are sufficient data for the purpose of this preliminary study and have allowed us to compare the performance of both model approaches. The ANN model used in the study could be successfully trained with the selected data subset size. In all trials, training and validation mean squared errors decreased monotonically to reasonably low minimum values. Besides, comparative analysis shows a consistently better performance of the data-driven model compared to the analytical model. Nevertheless, we acknowledge that increasing the size of the dataset with additional experimental data using multi-layered and heterogeneous tissue phantoms that better approximate biological tissue properties will improve the applicability of this model and this is planned as the next step in our research.

The experimental data used in this study were generated from insertion tests into transparent homogeneous soft tissue phantoms with and without addition of an artificial skin layer. The transparency of the material used to build the tissue samples allows the use of optical tracking as the ground truth method to measure needle deflection. However, this method is not applicable to testing with real biological soft tissue. The use of transparent phantom tissue and optical tracking with video cameras is a valid method to compare the prediction performance of the two proposed modelling approaches with the advantage that it allows performing multiple needle insertion tests in a controlled laboratory environment with a simple test setup and straightforward identification of the needle tip on the images. Nonetheless, our ultimate goal is to investigate the clinical application of the proposed ANN model and for that, it is necessary to train and test the model with data generated from needle insertion tests into representative biological soft tissue. Training and testing of the ANN's model performance for real heterogeneous biological tissue is planned as a future research work. To do this, optical tracking shall be replaced by US image tracking similarly to reference [13]. Ground truth needle deflection would be measured on US images after applying image processing algorithms to identify the location of the needle tip. 3D tracking of needle insertion would require an experimental setup with two US probes moving synchronized and parallel to the needle in order to track the position of the needle tip on two perpendicular planes.

IV. Conclusion and Future Work

Comparative performance analysis of the proposed analytical model based on beam deflection theory and the data-driven model using a feedforward ANN indicates a more reliable performance of the data-driven approach, which is also more robust to changes in the test conditions like tissue hardness or needle stiffness. Evaluation of the models' performance for different test data subsets resulted in accurate predictions of the proposed ANN architecture in 3 out of 4 testing scenarios, whereas the analytical model had larger estimation errors and its deflection predictions were not statistically comparable to ground truth data in any of the tested scenarios.

The findings of this study bring to light the potential of ANN data-driven models to be applied in clinical settings for imageless tracking of needle insertion procedures, in particular for brachytherapy needles. All analytical modelling approaches are built on assumptions that simplify the behaviour of the needle and surrounding tissue but which make these models unsuitable for application to a clinical setting due to the complex, anisotropic non-linear nature and huge variability of biological tissue properties. The proposed ANN model is capable of predicting 3D needle tip deflection with a submillimetre accuracy just from 9 input parameters. All variable input parameters can be measured in real time with a force sensor during needle insertion at small computational cost.

In order to confirm the applicability of the ANN approach to a clinical setting, further experimental testing will be carried out in order to train and test the ANN model with more representative data of biological tissue behaviour. Additional testing will be performed in multi-layered and heterogeneous tissue phantoms using the same ground truth method used in this study. Furthermore, experimental data will be obtained from needle insertion tests in more realistic heterogeneous biological tissue under controlled laboratory conditions and using US scanning as the ground truth method. Finally, a full-scale clinical testing is planned to obtain needle insertion data from multiple brachytherapy surgical procedures using TRUS imaging as a final validation step to confirm the suitability of the proposed ANN model for needle deflection prediction in the clinical setting.

References

[1] N. Abolhassani, R. V. Patel, and F. Ayazi, "Minimization of needle deflection in robot-assisted percutaneous therapy," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 3, no. 2, pp. 140–148, Jun. 2007.

[2] C. Avila Carrasco, "Modelling of brachytherapy needle deflection," *Master Diss. - Univ. West Engl.* September, 2019.

[3] C. M. Bishop, "Neural Networks," in Pattern Recognition and Machine Learning, New York, NY, USA: Springer, 2006, pp.225-290.

[4] M. W. T. Chao, P. Grimm, J. Yaxley, R. Jagavkar, M. Ng, and N. Lawrentschuk, “Brachytherapy: state-of-the-art radiotherapy in prostate cancer,” BJU Int., vol. 116, pp. 80–88, 2015.
[5] R. A. Cormack, C. M. Tempany, and A. V. D’Amico, “Optimizing target coverage by dosimetric feedback during prostate brachytherapy,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 48, no. 4, pp. 1245–1249, Nov. 2000.
[6] G. Fichtinger *et al.*, “Robotic assistance for ultrasound guided prostate brachytherapy,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2007, vol. 4791 LNCS, no. PART 1, pp. 119–127.
[7] J. M. Gere and B. J. Goodno, *Mechanics of Materials, 7th Edition*, Cengage Learning, 2009.
[8] D. Glozman and M. Shoham, “Image-guided robotic flexible needle steering,” *IEEE Trans. Robot.*, vol. 23, no. 3, pp. 459–467, Jun. 2007.
[9] O. Goksel, E. Dehghan, and S. E. Salcudean, “Modeling and simulation of flexible needles,” *Med. Eng. Phys.*, vol. 31, no. 9, pp. 1069–1078, Nov. 2009.
[10] M. Hudson Beale, M. T. B.Hagan, and H. B. and Demuth, “Neural Network Toolbox^{TM} User’s Guide R2017a,” *MathWorks*, p. 446, 2017.
[11] M. F. Jamaluddin *et al.*, “Quantifying 125 I placement accuracy in prostate brachytherapy using postimplant transrectal ultrasound images,” *Brachytherapy*, vol. 16, no. 2, pp. 306–312, Mar. 2017.
[12] H. Kataoka, T. Washio, M. Audette, and K. Mizuhara, “A model for relations between needle deflection, force, and thickness on needle penetration,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2001, vol. 2208, pp. 966–974.
[13] T. Lehmann, C. Rossa, N. Usmani, R. S. Sloboda, and M. Tavakoli, “A Real-Time Estimator for Needle Deflection During Insertion Into Soft Tissue Based on Adaptive Modeling of Needle-Tissue Interactions,” *IEEE/ASME Trans. Mechatronics*, vol. 21, no. 6, pp. 2601–2612, Dec. 2016.
[14] T. Lehmann, C. Rossa, N. Usmani, R. Sloboda, and M. Tavakoli, “A virtual sensor for needle deflection estimation during soft-tissue needle insertion,” in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 1217–1222.
[15] T. Lehmann, M. Tavakoli, N. Usmani, and R. Sloboda, “Force-sensor-based estimation of needle tip deflection in brachytherapy,” *J. Sensors*, vol. 2013, 2013.
[16] C. Mathworks, “Image Acquisition Toolbox ^{TM} User’s Guide R 2019a,” *MathWorks*, 2019.
[17] C. Mathworks, “Image Processing Toolbox ^{TM} User’s Guide R 2019a,” *MathWorks*, 2019.
[18] S. Misra, K. B. Reed, B. W. Schafer, K. T. Ramesh, and A. M. Okamura, “Mechanics of flexible needles robotically steered through soft tissue,” *Int. J. Rob. Res.*, vol. 29, no. 13, pp. 1640–1660, Nov. 2010.
[19] T. K. Podder *et al.*, “Effects of tip geometry of surgical needles: an assessment of force and deflection,” *IFMBE Proc*, vol. 11, no. 1, pp. 1727–1983, 2005.
[20] J. Poder, M. Carrara, A. Howie, D. Cutajar, J. Bucci, and A. Rosenfeld, “Derivation of in vivo source tracking error thresholds for TRUS-based HDR prostate brachytherapy through simulation of source positioning errors,” *Brachytherapy*, vol. 18, no. 5, pp. 711–719, Sep. 2019.
[21] T. Popescu, A. C. Kacsó, D. Pisla, and A. P. G. Kacsó, “Brachytherapy next generation: robotic systems,” *J. Contemp. Brachytherapy*, vol. 7, no. 6, p. 510, 2015.
[22] C. Rossa, T. Lehmann, R. Sloboda, N. Usmani, and M. Tavakoli, “A data-driven soft sensor for needle deflection in heterogeneous tissue using just-in-time modelling,” *Med. Biol. Eng. Comput.*, vol. 55, no. 8, pp. 1401–1414, Aug. 2017.
[23] H. Sadjadi, K. Hashtrudi-Zaad, and G. Fichtinger, “Needle deflection estimation: prostate brachytherapy phantom experiments,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 9, no. 6, pp. 921–929, 2014.
[24] Y. R. J. Van Veen, A. Jahya, and S. Misra, “Macroscopic and microscopic observations of needle insertion into gels,” *Proc. Inst. Mech. Eng. Part H J. Eng. Med.*, vol. 226, no. 6, pp. 441–449, Jun. 2012.
[25] R. J. Webster, J. S. Kim, N. J. Cowan, G. S. Chirikjian, and A. M. Okamura, “Nonholonomic Modeling of Needle Steering,” *Int. J. Rob. Res.*, vol. 25, no. 5–6, pp. 509–525, May 2006.

Description of the Video files added as supplemental material:

Ground truth needle deflections (x=vertical and y=horizontal) were measured through image processing of video-frames recorded by 2 cameras placed perpendicular to each other. One camera recorded the insertion process in the vertical plane (XZ) and the other camera recorded in the horizontal plane (YZ). The supplemental video files show examples of the videos recorded by both cameras during needle insertion tests.

1. Video_Plastic needle_Ver Plane.mp4

Shows a plastic needle insertion viewed from the camera recording in the vertical plane (XZ).

2. Video_Plastic needle_Hor Plane.mp4

Shows a plastic needle insertion viewed from the camera recording in the horizontal plane (YZ).

3. Video_Ti needle_Ver Plane.mp4

Shows a titanium needle insertion viewed from the camera recording in the vertical plane (XZ).

4. Video_Ti needle_Hor Plane.mp4

Shows a titanium needle insertion viewed from the camera recording in the horizontal plane (YZ).

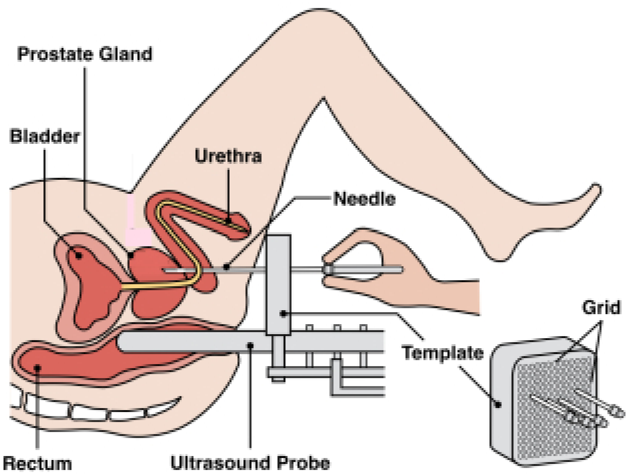


Fig. 1. Brachytherapy needle insertion procedure for prostate cancer treatment. In LDR brachytherapy radioactive seeds are permanently implanted in the prostate gland using needles inserted through a guiding grid template. The procedure is image-guided by transrectal ultrasound (TRUS).

115x86mm (72 x 72 DPI)

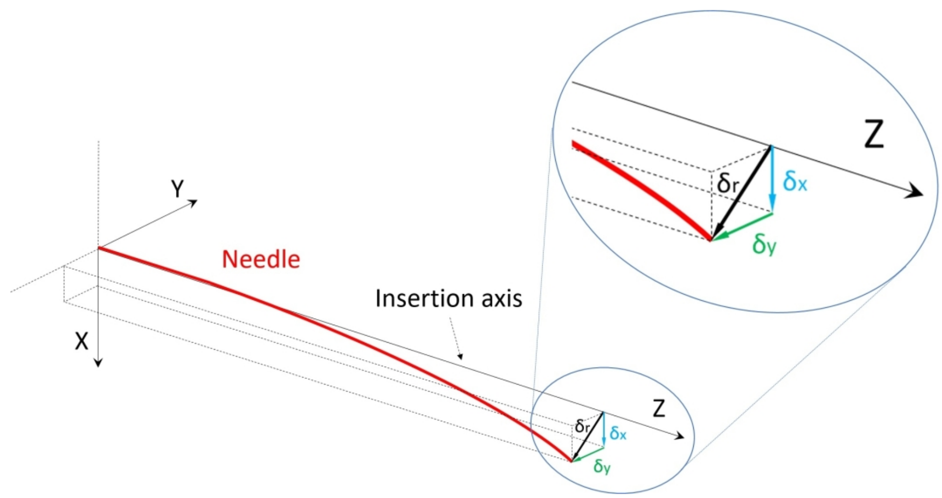


Fig. 2. Graphical representation of needle deflection. Total needle deflection in our study δr is defined as the Euclidean distance of the needle tip from the needle insertion axis; it is the resultant of perpendicular deflections δx (vertical) and δy (horizontal).

559x271mm (96 x 96 DPI)

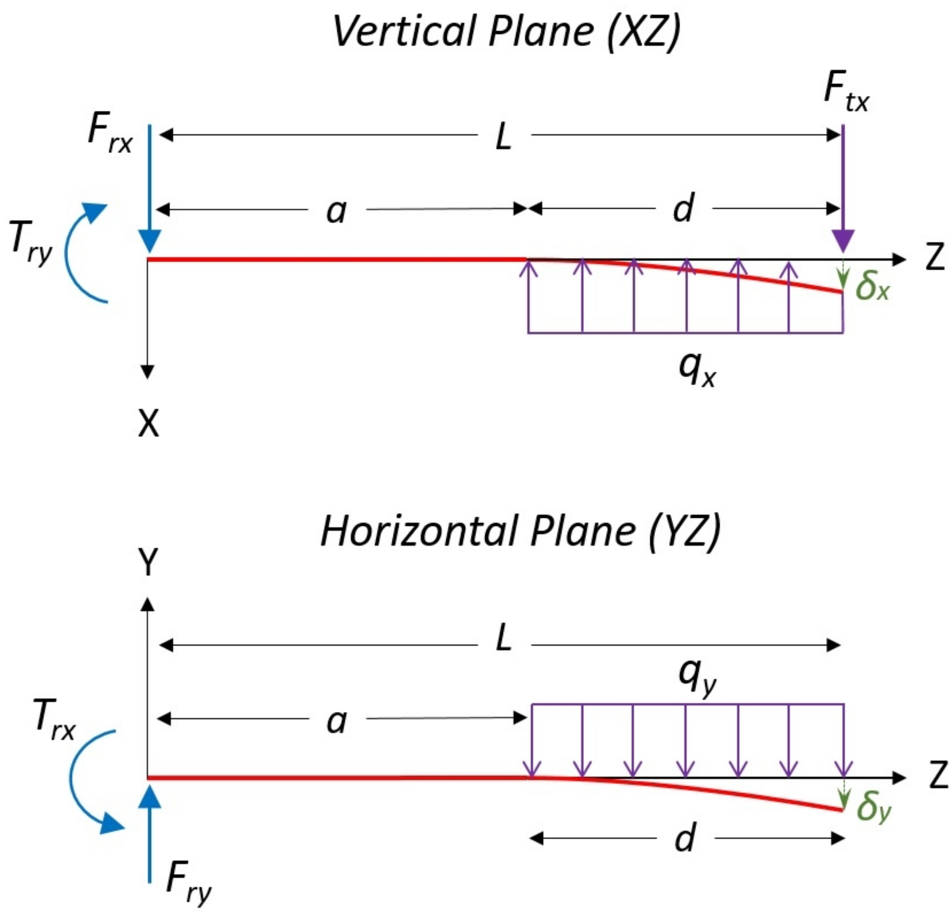


Fig. 3. Load profiles characterising the proposed beam-theory model. Top: Forces acting on the vertical plane (XZ); Bottom: Forces acting on the horizontal plane (YZ).

230x218mm (96 x 96 DPI)

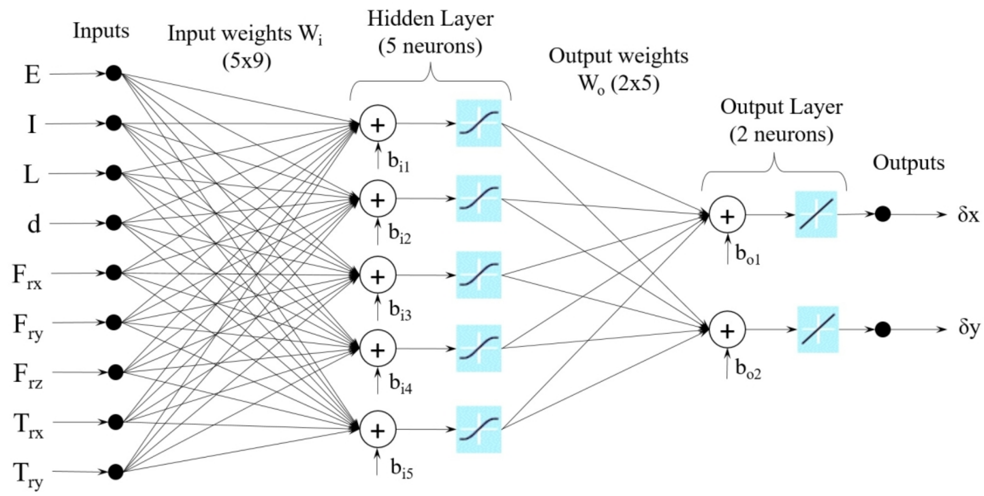


Fig. 4. Architecture of the proposed multilayer perceptron feed-forward artificial neural network for the data-driven model. The network has one hidden layer with 5 neurons using sigmoid functions and one output layer with 2 linear neurons.

482x245mm (96 x 96 DPI)

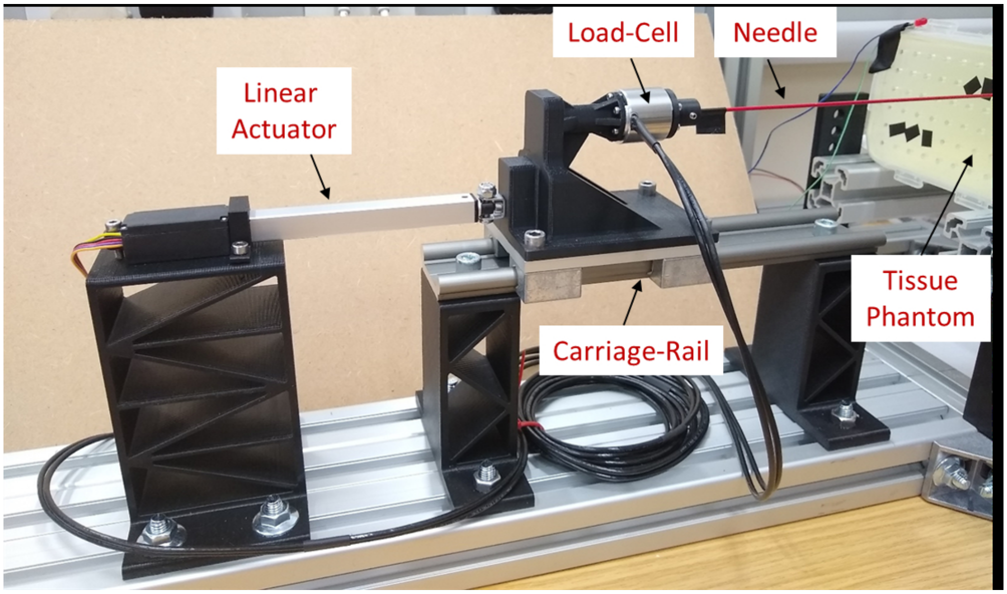


Fig 5.a. Experimental set-up for needle insertion tests. Top image: Mechatronic device for needle insertion and retraction using a linear actuator attached to a carriage holding the needle.

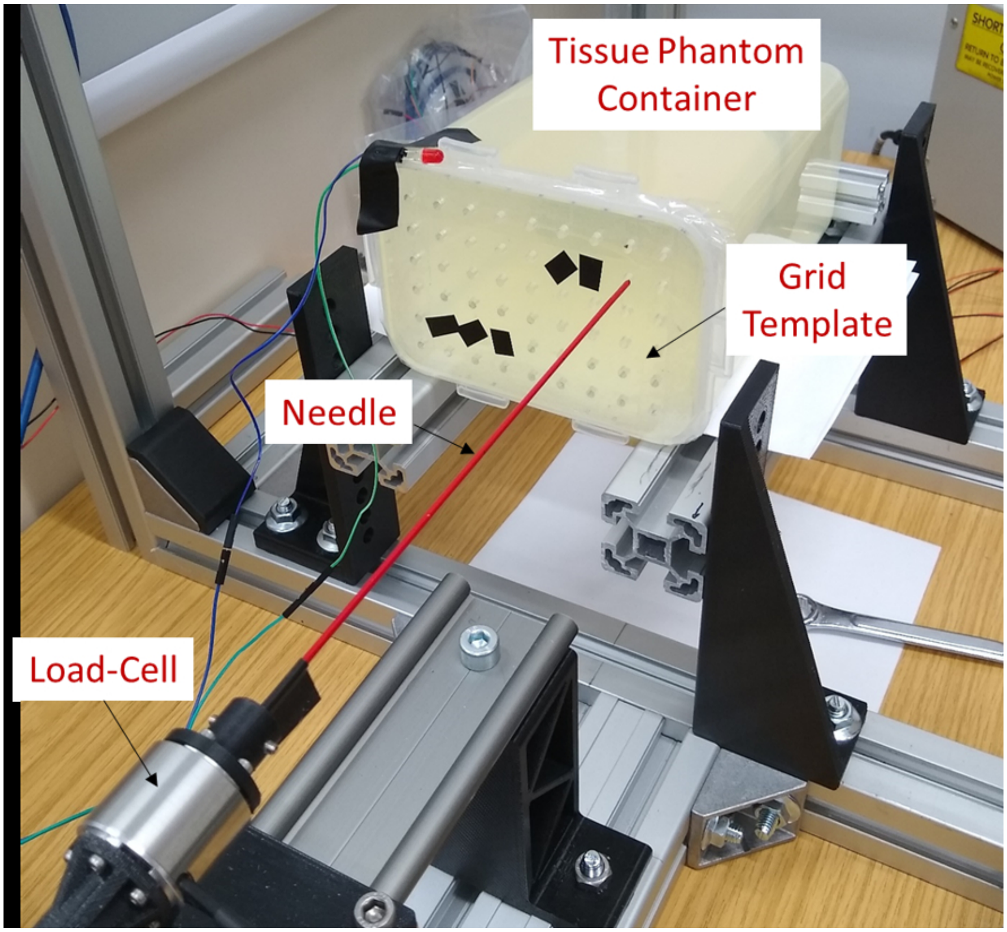


Fig 5.b. Experimental set-up for needle insertion tests. Bottom image: Detail of the needle mounted on the insertion device and tissue phantom container. A load-cell is fixed to the base of the needle holder and measures reaction forces and torques at the base of the needle.

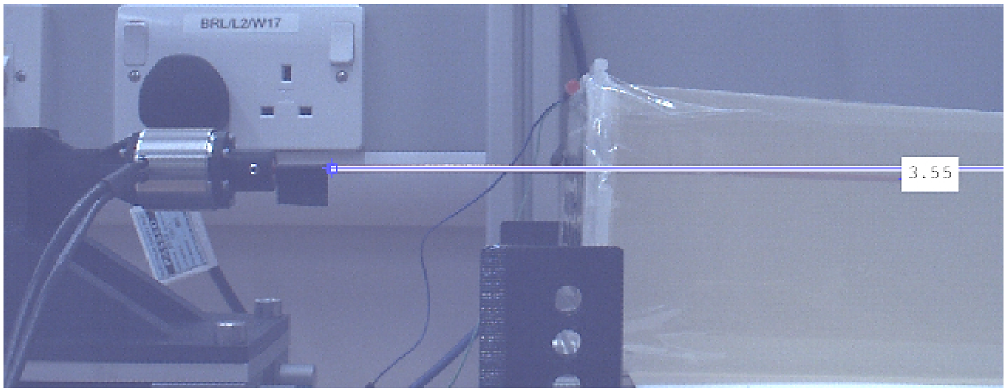


Fig. 6.a. Example of ground truth deflection measurement. Tip deflection values in vertical and horizontal planes for a certain insertion depth are obtained from image processing of the corresponding video-frames. Top image corresponds to the camera for the vertical plane (XZ) and bottom image corresponds to the horizontal plane camera (YZ).

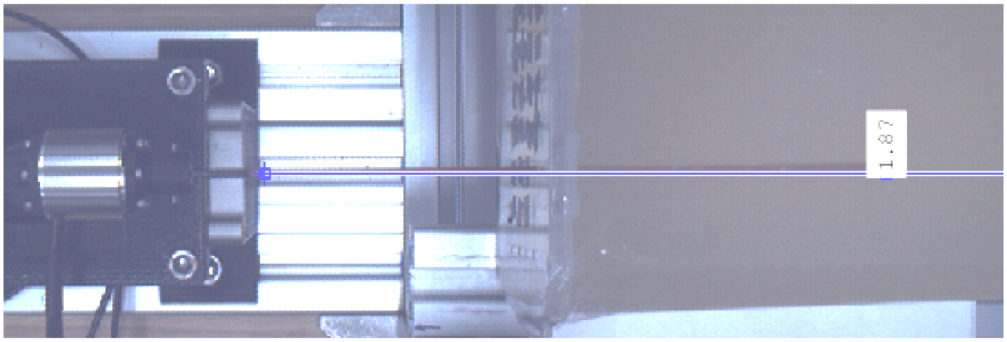


Fig. 6.b. Example of ground truth deflection measurement. Tip deflection values in vertical and horizontal planes for a certain insertion depth are obtained from image processing of the corresponding video-frames. Top image corresponds to the camera for the vertical plane (XZ) and bottom image corresponds to the horizontal plane camera (YZ).

Fig. 7. Detail of the tip geometry of the brachytherapy needles used in the study. Left: Conical tip plastic needle with 2mm diameter. Right: Sharp trocar tip titanium needle with 1.65mm diameter.

280x83mm (96 x 96 DPI)

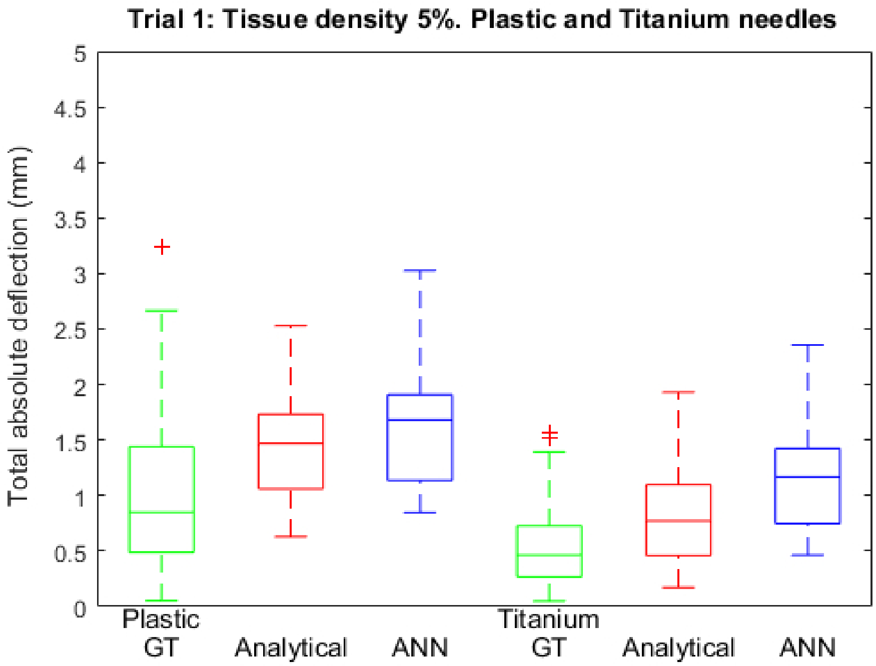


Fig 8.a. Box-plot distributions of ground truth (GT) and predicted (Analytical & ANN) total absolute needle tip deflection (|δr|) grouped by needle type (Plastic / Titanium) and soft tissue hardness (5, 10 and 20% densities). Individual plots show the absolute deflections for each of the test data subsets defined in Table I (Trial 1 (a), Trial 2 (b), Trial 3 (c) and Trial 4 (d)). Box-plots include data between the first and third quartiles, whiskers extend up to ±2.7σ and red crosses represent outliers. Box colours are used for better clarity to distinguish GT (Green), Analytical (Red) and ANN (Blue) deflections.

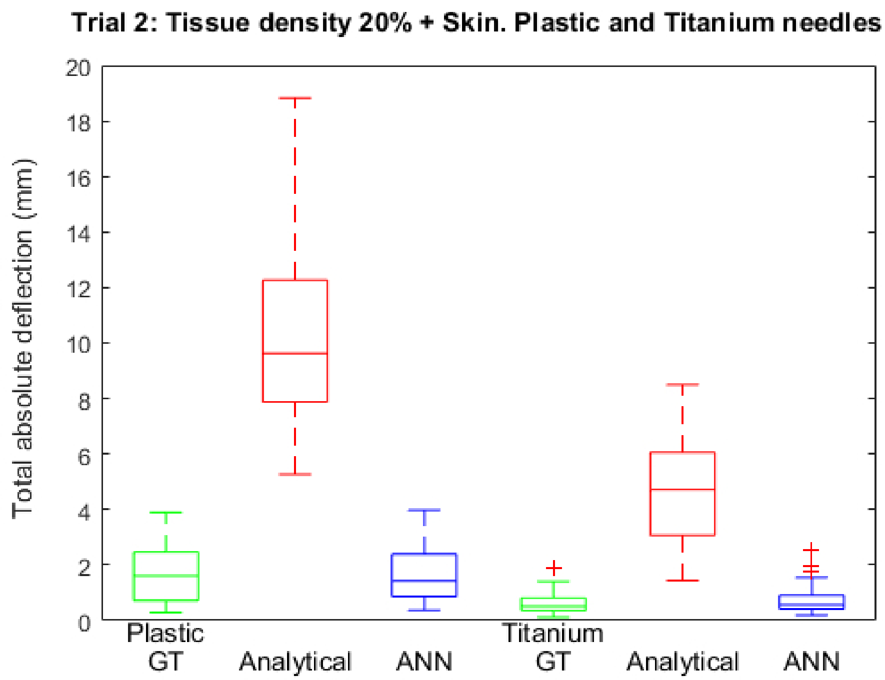


Fig 8.b. Box-plot distributions of ground truth (GT) and predicted (Analytical & ANN) total absolute needle tip deflection (|δr|) grouped by needle type (Plastic / Titanium) and soft tissue hardness (5, 10 and 20% densities). Individual plots show the absolute deflections for each of the test data subsets defined in Table I (Trial 1 (a), Trial 2 (b), Trial 3 (c) and Trial 4 (d)). Box-plots include data between the first and third quartiles, whiskers extend up to ±2.7σ and red crosses represent outliers. Box colours are used for better clarity to distinguish GT (Green), Analytical (Red) and ANN (Blue) deflections.

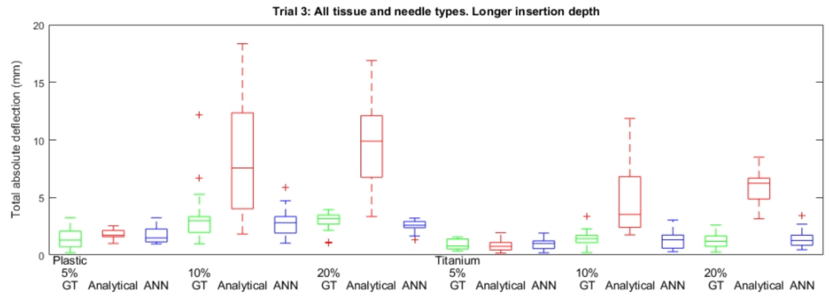


Fig 8.c. Box-plot distributions of ground truth (GT) and predicted (Analytical & ANN) total absolute needle tip deflection (|δr|) grouped by needle type (Plastic / Titanium) and soft tissue hardness (5, 10 and 20% densities). Individual plots show the absolute deflections for each of the test data subsets defined in Table I (Trial 1 (a), Trial 2 (b), Trial 3 (c) and Trial 4 (d)). Box-plots include data between the first and third quartiles, whiskers extend up to ±2.7σ and red crosses represent outliers. Box colours are used for better clarity to distinguish GT (Green), Analytical (Red) and ANN (Blue) deflections.

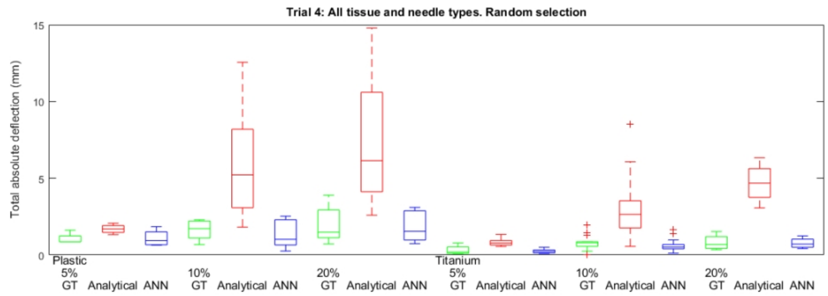


Fig 8.d. Box-plot distributions of ground truth (GT) and predicted (Analytical & ANN) total absolute needle tip deflection (|δr|) grouped by needle type (Plastic / Titanium) and soft tissue hardness (5, 10 and 20% densities). Individual plots show the absolute deflections for each of the test data subsets defined in Table I (Trial 1 (a), Trial 2 (b), Trial 3 (c) and Trial 4 (d)). Box-plots include data between the first and third quartiles, whiskers extend up to ±2.7σ and red crosses represent outliers. Box colours are used for better clarity to distinguish GT (Green), Analytical (Red) and ANN (Blue) deflections.

TABLE I
DATA DISTRIBUTION FOR THE 4 ANN MODEL VARIANTS GENERATED IN THE STUDY (CHARACTERISTICS AND SAMPLE SIZE OF EACH DATA SUBSET)

ANN Model	Training Data		Validation Data		Test Data	
	Description	Data points	Description	Data points	Description	Data points
Trial 1	10% and 20% Tissue Density	369	10% and 20% Tissue Density	41	5% Tissue Density	80
Trial 2	5% and 10% Tissue Density + 20% Density with no skin	369	5% and 10% Tissue Density + 20% Density with no skin	41	20% Tissue Density + Skin	80
Trial 3	Insertion depths ≤ 75mm	324	Insertion depths ≤ 75mm	36	Insertion depths = 97mm	123
Trial 4	Random	394	Random	49	Random	49

TABLE II

SUMMARY OF NEEDLE INSERTION TESTS PERFORMED IN THE STUDY GROUPED PER COMBINATIONS OF NEEDLE TYPE, TISSUE DENSITY AND USE OF ADDITIONAL SKIN TISSUE LAYER. FOR EACH GROUP, TESTS WERE EQUALLY SPLIT INTO TWO AVERAGE INSERTION SPEEDS OF 10MM/S AND 15MM/S

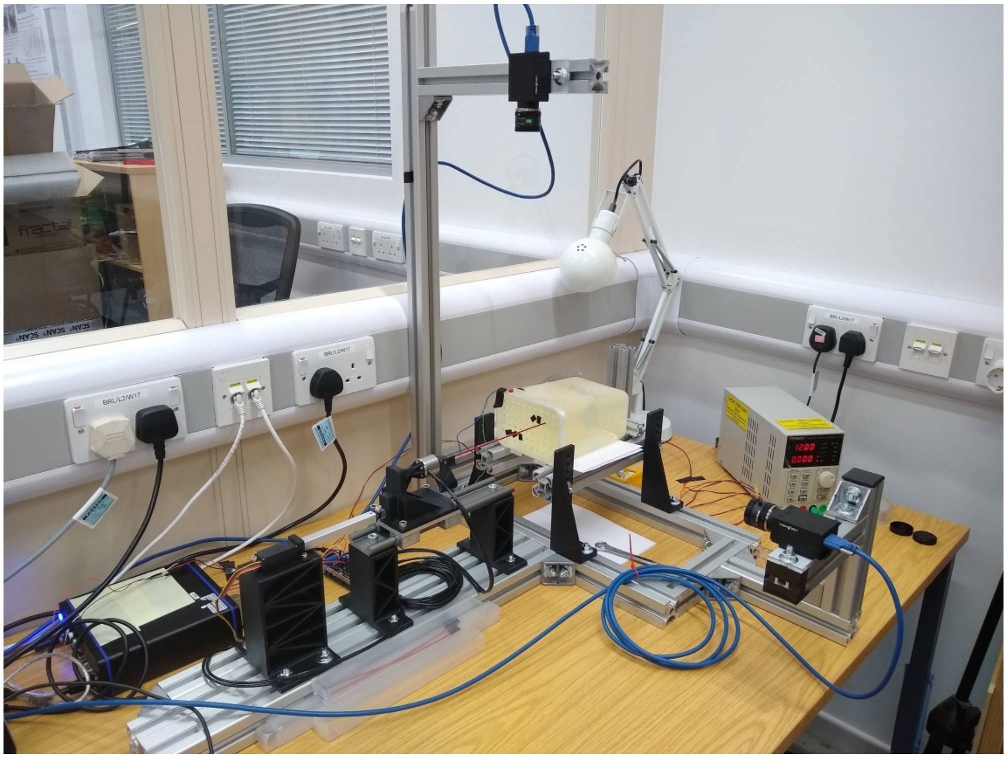
Needle Type	Tissue Density	Skin Layer	Number of Tests
Plastic	5%	No	10
	10%	No	9
		Yes	18
	20%	No	10
		Yes	10
Titanium	5%	No	10
	10%	No	20
		Yes	16
	20%	No	10
		Yes	10

TABLE III

COMPARISON OF ANALYTICAL AND ANN MODELS PERFORMANCE FOR PREDICTION OF TOTAL ABSOLUTE TIP DEFLECTIONS IN FOUR DIFFERENT TEST DATA SUBSETS

Results for Total Absolute Needle Tip Deflection ($\|\delta r\|$)					
	Average Deflection ± STD (mm)			MAE ± STD (mm)	
Test Data Subset	Ground Truth	Analytical Model [a]	ANN Model [a]	Analytical Model	ANN Model
Trial 1	0.78 ± 0.61	1.14 ± 0.55 *	1.37 ± 0.57 *	0.50 ± 0.29	0.62 ± 0.41
Trial 2	1.15 ± 0.93	7.49 ± 4.15 *	1.19 ± 0.88	6.35 ± 3.47	0.32 ± 0.28
Trial 3	2.00 ± 1.46	6.13 ± 4.65 *	1.84 ± 1.03	4.23 ± 3.82	0.50 ± 0.82
Trial 4	1.13 ± 0.84	4.13 ± 3.51 *	0.99 ± 0.81	3.03 ± 3.04	0.29 ± 0.22

[a] * Hypothesis of equality of means between ground truth and predicted deflections is rejected (*Two-sample t-tests*)



1444x1083mm (72 x 72 DPI)