Real-Time Visualization and Analysis of Clinicians' Performance During Palpation in Physical Examinations

Alejandro Granados, Naomi Low-Beer, Jenny Higham, Roger L. Kneebone, and Fernando Bello

***Abstract*—*Objective:* Motivated by the fact that palpation skills are challenging to learn and teach, particularly during Digital Rectal Examinations (DRE), and the lack of understanding of what constitutes adequate performance, we present a visualization and analysis system that uses small position and pressure sensors located on the examining finger, allowing the quantitative analysis of duration, steps, and pressure applied. *Methods:* The system is first described, followed by an experimental study of 20 experts from four clinical specialties performing ten DREs each on a benchtop model using the proposed system. Analysis of the constitutive steps was conducted to improve understanding of the examination. A Markov model representing executed tasks and analysis of pressure applied is also introduced. *Results:* The proposed system successfully allowed the visualization and analysis during the experimental study. General practitioners and nurses were found to execute more tasks compared to urologists and colorectal surgeons. Urologists executed the least number of tasks and were the most consistent group compared to others. *Conclusion:* The ability to "see through" allowed us to better characterize the performance of experts when conducting a DRE on a benchtop model, comparing the performance of relevant specialties, and studying executed tasks and the pressure applied. The Markov model presented summarizes task execution of experts and could be used to compare the performance of novices against that of experts. *Significance:* This approach allows for the analysis of performance based on continuous sensor data recording that can be easily extended to real subjects and other types of physical examinations.**

***Index Terms*—Digital rectal examination, sensors, palpation, quantitative analysis, performance.**

Manuscript received June 26, 2017; revised October 23, 2017; accepted November 24, 2017. Date of publication March 28, 2018; date of current version August 20, 2018. This work was supported in part by grants from the EPSRC, the U.K. Engineering and Physical Science Research Council, the London Deanery STeLI initiative, and the Health Education North West London. *(Corresponding author: Fernando Bello.)*

A. Granados and F. Bello are with the Simulation and Modelling in Medicine and Surgery Group, Centre for Engagement and Simulation Science, Chelsea & Westminster Hospital, Imperial College London, London SW7 2AZ, U.K. (e-mail: a.granados@imperial.ac.uk; f.bello@imperial.ac.uk).

N. Low-Beer was with Imperial College London, London SW7 2AZ, U.K. She is now with the Lee Kong Chian School of Medicine, Singapore.

J. Higham was with Imperial College London, London SW7 2AZ, U.K. She is now with St. George's Hospital, University of London, London WC1E 7HU, U.K.

R. L. Kneebone is with the Centre for Engagement and Simulation Science, Chelsea & Westminster Hospital, Imperial College London, London SW7 2AZ, U.K.

Digital Object Identifier 10.1109/TBME.2017.2780982

I. Introduction

PALPATION, considered an open-ended task that seeks for an undefined solution, requires different types of knowledge, motor and perceptual skills, as well as a therapeutic attitude [1]. Palpation is routinely performed before and during medical procedures and is fundamental during a physical examination [2]. Clinicians use their finger(s) on patients to detect landmarks in the body, assess physiology of systems and diagnose abnormalities that either occur on the skin, under the skin, inside an organ or in areas that are reachable, but impossible to see. They do so by determining size, shape, location, texture, temperature, mobility, pulsation, fluctuation, guarding and tenderness [3], using exploratory procedures including lateral motion, pressure, static contact, unsupported holding, enclosure and contour following [4]. During this process, the spatial distribution of forces at the fingertip, kinaesthetic feedback, as well as the duration of the exploration, contribute to the generation of tactile information [4]. Palpation skills are difficult to learn since they require a highly trained sense of touch, patient variability is inherent and opportunities for practice are limited and hard to standardise [5]. Although simulation offers obvious benefits to overcome some of these limitations, palpation is mostly neglected in medical training simulators [2]. A possible reason of this is that palpation skills are challenging to teach and assess [1], given the limitations of articulating how objects should be perceived, in addition to highlighting diagnosis as a means to evaluate performance, rather than palpation techniques that ensue competence.

The intimate nature of some palpation-based examinations along with the inability to see what is palpated in constrained spaces, imposes additional limitations on teaching and learning. An example of these examinations includes Digital Rectal Examinations (DREs), where a clinician inserts the index finger through the anus of a patient to diagnose anorectal [6], [7] and prostate abnormalities [8]. A crucial problem is that there is a lack of understanding of what are the pressure and palpation techniques that lead to an adequate examination. Previous attempts have studied DRE qualitatively and have reported a series of steps in the form of a Cognitive Tasks Analysis (CTA) based on direct observation of finger movement on a standard DRE

(a) (b)

Fig. 1. (a) A Micro 6DOF position sensor coil on the nail of the index finger before covering it with a capacitive pressure sensor for DRE. (b) Set up with electromagnetic tracker behind the DRE benchtop model.

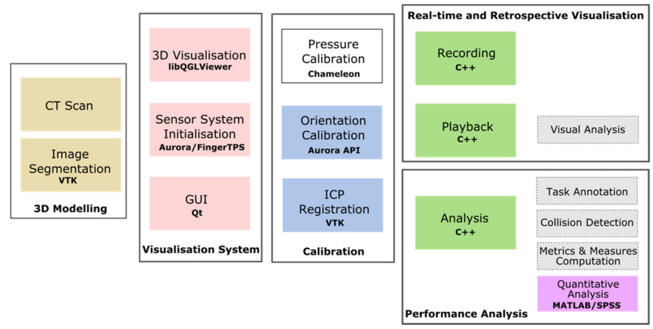


Fig. 2. Components of the proposed 3D visualisation and analysis system for palpation skills. 3D surface models of a benchtop model are generated from CT scans and then rendered transparently in 3D. Position and pressure sensors are integrated, initialised and calibrated. Our system allows for recording, playback and analysis.

part-task trainer benchtop model with the rectum cut away [9]. Quantitative attempts have focused on computing performance metrics from pressure sensors embedded on an instrumented prostate benchtop model [10]–[12]. However, by using a discrete number of sensors on fixed anatomical locations, the proposed systems not only fail to capture other important regions such as the rectal walls, but are also unable to offer a continuous pressure map across the anatomy to be examined. Better understanding of palpation techniques on internal anatomy throughout the rectal examination may allow us to assess which techniques lead to competent DRE performance.

In this paper, we present a visualisation and analysis system using small position and pressure sensors that are located on the examining finger, allowing the quantitative analysis of performance by expert clinicians that routinely perform this examination (General Practitioners, Nurse Practitioners, Urologists and Colorectal surgeons), including duration, sequence of steps and pressure applied. Our primary goal is to improve our understanding of palpation during a physical examination, in this case DRE, through the further characterisation and analysis of its constitutive steps by using position and pressure sensor technology on a plastic part-task trainer benchtop model. First, the sensor technology, DRE model set up and visualisation components are described, followed by the analysis methodology and experimental study. Results of questionnaires and analysis of tasks and pressure applied during internal examinations are then presented, followed by a discussion and conclusions of the visualisation and analysis system, methodology and findings.

II. Materials and Methods

A. Position Tracking and Pressure Sensing

Before donning a clinical glove, a position sensor coil (Aurora Micro 6DOF 0.8 mm × 9 mm) was placed on the nail of the examining finger and tracked with an electromagnetic tracker (NDI Aurora, tracking volume 50 × 50 × 50 mm) located behind a Limbs & Things Rectal Examination Trainer Mk1 part-task trainer benchtop model. A capacitive pressure sensor pad (Pressure Profile System FingerTPS) located on the pad of the index finger was used to capture pressure during the examination (Fig. 1).

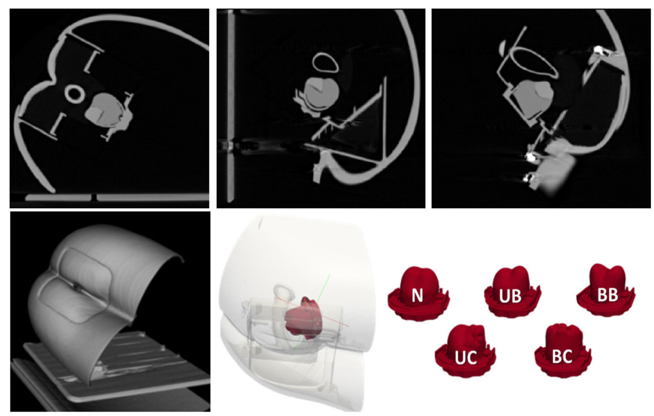


Fig. 3. Sample scan planes of the DRE benchtop model CT (above) and segmented 3D models of skin and prostate (below): Normal (N), Unilateral Benign (UB), Bilateral Benign (BB), Unilateral Carcinoma (UC) and Bilateral Carcinoma (BC).

B. 3D Real-Time Visualisation

The visualisation and analysis system is based in the framework introduced in [13]. It allows real-time visualisation during recording, as well as retrospective playback. It also permits continuous data recording (40 Hz sampling rate) of examining finger position, orientation and pressure while palpating any internal structure during the examination (Fig. 2). The proposed system was developed in C++ using libQGLViewer for graphic rendering, Qt for Graphical User Interface (GUI). The Aurora NDI SDK and FingerTPS API were integrated into a multi-threading system.

3D models of the benchtop model, together with the available five different prostate types (Normal (N), Unilateral Benign (UB), Bilateral Benign (BB), Unilateral Carcinoma (UC) and Bilateral Carcinoma (BC)), were generated from a high resolution CT scan by constructing surface models using marching cubes in VTK (Fig. 3).

The calibration process consists of three stages that take place before performing the examination, once the position and pressure sensors are placed on the examining finger and the hand is covered by the clinical glove. Firstly, the pressure sensor is

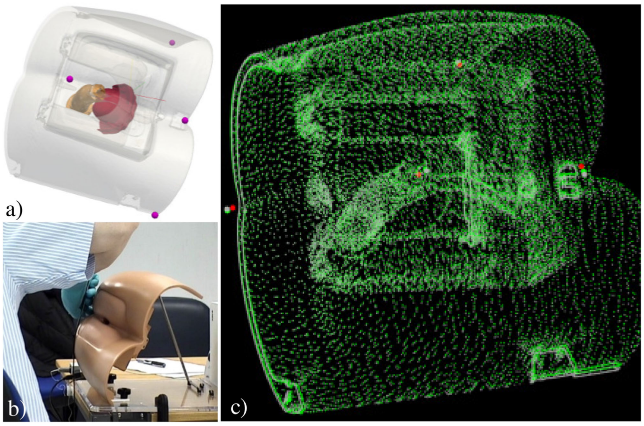


Fig. 4. Registration process: (a) four landmarks used for registration, (b) an example of a participant touching the fourth landmark, (c) visual result of registration comparing original (grey-coloured) and registered (green-coloured) mesh with position of landmarks (red points).

calibrated using the FingerTPS Chameleon software using a reference sensor, a process resulting in updated calibration system files within our system. Then, the participant is asked to point their tracked index finger towards the electromagnetic tracker to calibrate the location of the position sensor on the finger. Lastly, four anatomical landmarks are touched by the participant using the tracked index finger. These landmarks are used to register the 3D surface models with the corresponding benchtop model set up using the standard Iterative Closest Point (ICP) algorithm in VTK (Fig. 4). A 4×4 registration matrix is computed as a result of the ICP, parsed and saved by the system.

3D surface models are then loaded, registered and plotted semi-transparently in the visualisation and analysis system. A 3D mesh representation of the examining finger is translated and rotated according to the position sensor, as well as colour-coded to indicate the amount of pressure recorded by the relevant pressure sensor at that particular anatomical location. Two views are presented: a sagittal view (left) and a coronal view (right), along with a pressure plot (bottom) indicating the applied pressure at a particular point in time (Fig. 5).

C. Task Decomposition and Annotation

The CTA presented in [4] was used to annotate relevant steps observed whilst performing a DRE (Table I).

During playback, CTA steps were manually identified and labelled by selecting the start and end intervals of the observed tasks in the pressure plot pane of our visualisation system (Fig. 6). The annotated tasks allow us to study their duration, frequency, trajectories, forces and other metrics.

D. Performance Analysis

1) Tasks: The type of tasks performed, their duration, the sequence of task execution and the pressure applied during these tasks are reported as measures of performance.

2) Prostate: The 3D model of the prostate was divided into ten regions. The region palpated by the tracked index finger is reported through a collision detection mechanism based on an Axis Aligned Bounding (AABB) tree [13]. The number and frequency of regions palpated, the type of region (either normal, enlarged or abnormal), the orientation of the finger and the pressure applied are also reported and used to compute metrics and palpation primitives (Fig. 7). One of the most important metrics reported by clinicians is *completeness*, which indicates how much of the prostate gland (posterior and palpable regions) was palpated fully once.

3) Anal Canal and Rectum: The palpation of the anal canal and rectal walls is abstracted into a polar coordinate map consisting of regions based on finger orientation (with sectors every 60°) and finger insertion depth (concentric rings every 2 cm) [13]. Completeness is then computed based on the number of palpated regions. Palpation primitives are computed similarly to the prostate.

E. Experimental Study

The purpose of our experimental study was to improve our understanding of palpation during a DRE by quantifying performance of clinicians from different specialities that routinely perform this examination, through a task analysis which included: duration, sequence of steps and pressure applied.

A convenience sample of 20 experts from the following four different clinical disciplines was selected: General Practitioners (GP), Nurse Practitioners (NP), Urologists (UR) and Colorectal surgeons (CR). Each performed ten examinations on the DRE plastic part-task trainer benchtop model using either normal or abnormal (enlarged or carcinoma) prostate models (two examinations for each prostate type) (Fig. 3). Participants were asked to concentrate on the steps that need to be performed during DRE for the first five examinations, whereas they were asked to give a diagnosis only for the remaining 5 examinations ($N = 198$ with two examinations not recorded).

Demographics and an end-of-study structured questionnaire comprising items rated on a 5-point Likert scale ($1 =$ strongly disagree to $5 =$ strongly agree) were completed by participants to capture their experience with the sensors during the study. Ethics approval was granted by the NHS National Patient Safety Agency Research Ethics Committee (Reference number: 09/H0701/68).

A 1-Between (speciality) and 2-Within (prostate type & diagnosis) Linear Mixed Model (LMM) with maximum likelihood (ML) was used to report statistically significant differences. Total variability was studied using an intra-class correlation model ICC(2, 1) with participants as raters and prostate types as cases.

III. Results

A. 3D Visualisation and Analysis System

The proposed system successfully allowed the visualisation, recording and analysis during our experimental study. We were able to see through using different views and observe finger movement and orientation during the internal examination. Both the colour-mapped pressure and the pressure plot were useful to better understand palpation. Highlighting time intervals of executed tasks was simple for task annotation, which allowed

Fig. 5. Recording, visualisation and analysis system with two views of the anatomy and a pressure plot during DRE of a normal prostate.

TABLE I
RELEVANT TASKS (ADAPTED) FOR ANALYSIS BASED ON CTA

Task	DRE task
23	Position pad of right index finger on anus
24	Apply gentle pressure with finger pad on anus for a few seconds
26	Insert finger with pad posteriorly
27	Assessment of sphincter tone
28	Insert finger beyond sphincter into rectum
29	Coccyx is reached
32	Rectal wall palpation: start circumferential palpation at level of coccyx
33	Rectal wall palpation: systematic, full 360 degree sweep
34	Prostate palpation
45	Remove finger

Physical and sensorial tasks for the internal examination stage during DRE.

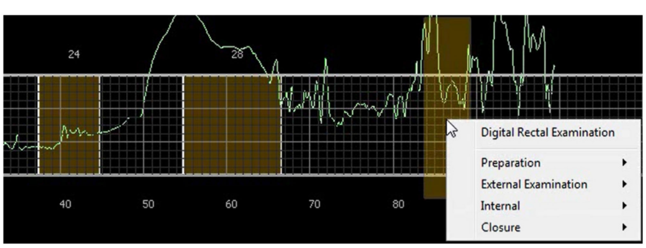


Fig. 6. Task annotation based on CTA. Pressure plot pane is used to select intervals (shaded brown areas) in order to label an observed task (number above intervals indicate task ID), particularly those tasks related to the internal examination stage that are possible to see through our system.

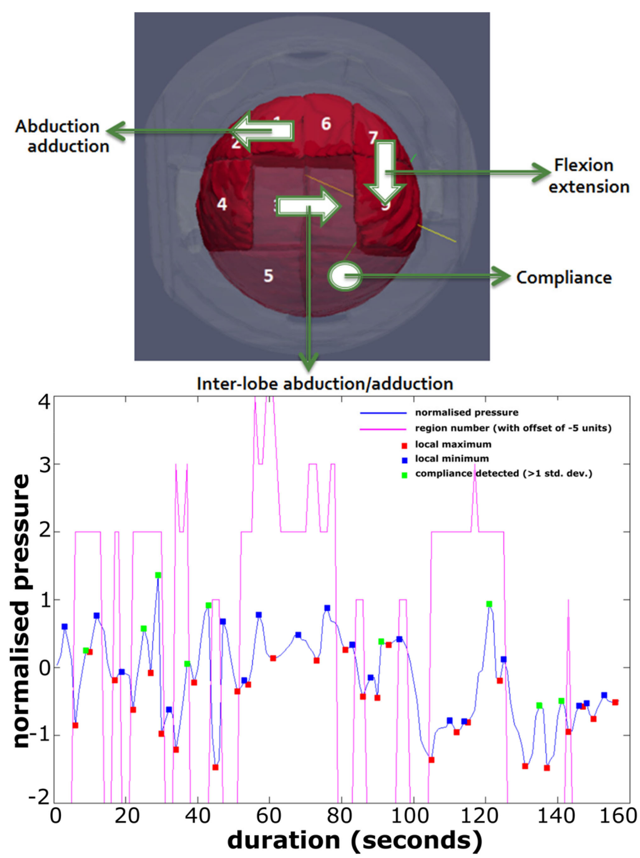


Fig. 7. **Top**: Palpation primitives on the prostate: abduction/adduction (lateral movement), flexion/extension (movement amongst the base, mid and apex sections of the prostate), supination/pronation (finger rotated $\pm 45°$) and compliance. **Bottom**: Z-score or standard-score normalised pressure plot during prostate palpation (blue series) showing how compliance (green squares) is computed based on the applied pressure in a single region (pink series) when the difference between local minima (red squares) and local maxima (blue squares) is greater than one standard deviation.

our system to generate measures and metrics related to executed tasks, as well as prostate and rectal walls palpation. The pressure calibration process within the FingerTPS Chameleon software varied amongst participants and pressure sensors were subject to drift. Pressure data was normalised using the Z-score or standard score (number of standard deviations away from the mean) to account for the variation in the calibration process, whereas drift was minimised by setting the baseline before recording

TABLE II
CONFUSION MATRIX OF DIAGNOSES

		Diagnosis by clinician					
		N	UB	BB	UC	BC	Sensitivity
Model	N	8	11	1			0.4
	UB	8	8	4			0.4
	BB	1	9	6	3	1	0.3
	UC	2	1		16	1	0.8
	BC		1	1	2	16	0.8

Type of prostate used for the examination (rows) and diagnosis made by participant after examination (columns). Cells represent distribution frequency (experiments). Well-diagnosed (true-positive) cases are located in the diagonal. Carcinoma cases that were misdiagnosed and benign cases misdiagnosed as normal are shown in red, whereas green indicates normal or benign cases misdiagnosed as carcinoma. *Types of prostate*: Normal (N), Unilateral Benign (UB), Bilateral Benign (BB), Unilateral Carcinoma (UC) and Bilateral Carcinoma (BC).

TABLE III
QUESTIONNAIRES

	Question	GP	NP	UR	CR	Group
Pressure and position sensors	The pressure and tracking sensors are unobtrusive	3.8 (1.3)	4.4 (.6)	3.8 (1.1)	3.6 (.6)	
	The pressure and tracking sensors could be used as a DRE teaching aid	4.8 (.5)	4.8 (.5)	4.8 (.5)	4.2 (.5)	**4 (1)**
	The pressure sensor was comfortable to wear and did NOT affected my examination and diagnostic skills	3.2 (1.3)	4 (1.2)	3.2 (1.1)	3.2 (.8)	

Means (standard deviation) of questionnaires by clinical discipline and overall (1-Strongly disagree; 2-Disagree; 3-Neither agree nor disagree; 4-Agree; 5-Strongly agree). *Clinical disciplines*: General Practitioner (GP), Nurse Practitioner (NP), Urologist (UR) and Colorectal surgeon (CR).

each examination and by following the manufacturer's recommendations, including: use correct size of pad on index finger and ensure pad is not loose after placing the clinical glove on the hand to maximise surface contact, allow light use for 4–5 min before calibration to minimise temperature effects, run calibration smoothly avoiding sudden changes in pressure applied, and gently wiggle finger whilst sensor is worn to discard unwanted changes in force magnitude.

B. Diagnosis and Questionnaires

The diagnosis by participants is presented in Table II. Values in the diagonal indicate correct diagnoses (sensitivity), whilst values off the diagonal indicate incorrect ones from a total of 100 diagnoses (five for each participant). It can be observed that abnormal prostate types (UC and BC) were easier to diagnose compared to other prostate types. Bilateral Benign prostate was the most misdiagnosed. The impact of misdiagnosis can be better understood by calculating sensitivity for benign (N, UB, and BB) and abnormal (UC and BC) cases, which results in values of 0.95 and 0.75, respectively.

The participants agreed that the sensors were unobtrusive, but they neither agreed nor disagreed that the pressure sensor was comfortable to wear and did not affected their performance. They strongly agreed that the sensors could be used as a teaching aid (M = 4.65, SD = 0.49) (Table III).

C. Task Analysis

1) Duration: Table IV shows a summary of the main findings related to the tasks performed during DRE, along with a basic stratification of the duration across specialties.

2) Sequence: It was observed that the rectum is palpated more than once after prostate palpation by NPs (1.56 times), sometimes by GPs (0.44 times) and CRs (0.58 times), but less frequently by URs (0.2 times) (Table V). The prostate was palpated nearly twice after rectal palpation by NPs (1.7 times), nearly once by GPs (0.88 times) and CRs (0.94 times) and rarely by URs (0.08 times), whereas the prostate was palpated nearly once after inserting the finger beyond the sphincters into the rectum by URs (0.94 times), sometimes by GPs (0.4 times), and even less frequently by NPs and CRs (0.26 times).

The finger is removed nearly once after prostate palpation by GPs (0.9 times), URs (0.82 times) and CRs (0.72 times) and less frequently by NPs (0.44 times), whereas the finger is removed sometimes after rectal palpation by NPs (0.54 times) and less frequently by CRs (0.26 times), GPs (0.1 times) and URs (0.02 times).

Following these observations, a Markov model was computed and is presented in Fig. 8. The transition probabilities are averaged across experiments and characterise an examination on a plastic *prostate-only* benchtop model. A DRE starts when participants position the pad of the index finger on the anus (task 23). After this, they rarely apply gentle pressure (task 24; probability = .16) and they either insert their finger with pad posteriorly (task 26; probability = .52) or anteriorly followed up by insertion of finger beyond sphincter into rectum (task 28; probability = .32), which is commonly done by urologists.

The assessment of sphincter tone is rarely executed after initial finger insertion (task 27; probability = .03) followed by further insertion of finger into rectum (task 28; probability = 1.0). Once within the rectal walls, three possibilities may occur: either a) the coccyx is palpated (task 29; probability = .36) most likely if the finger was inserted posteriorly, b) the rectum is palpated (task 33; p = .17), or c) the prostate is palpated (task 34; probability = .46). What occurs thereafter is a combination of these three tasks and can be better characterised by different clinical specialties (see Table V) before removing the examining finger (task 45).

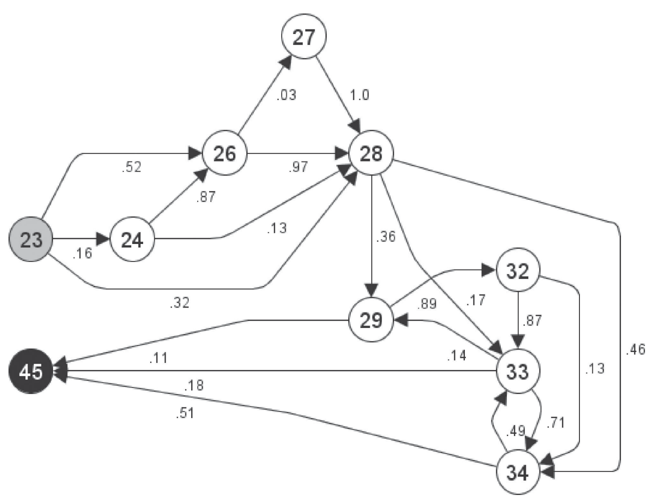
D. Pressure Analysis

Pressure data was normalised based on the z-score (number of standard deviations apart from the mean). Fig. 9 (top) shows an example of the force profile for one of the examinations. Normalised pressure data during prostate palpation (task 34) was transformed from the time domain to the frequency domain in MATLAB to obtain the most prominent frequency of applied

TABLE IV
SUMMARY OF FINDINGS RELATED TO THE DURATION OF TASKS (SECONDS)

Task	Mean	GP	NP	UR	CR	Notes
Apply gentle pressure (task 24)	1.9 (1.2)	2.2 (.3)	1.6 (.3)	-	-	Only one GP and one NP applied gentle pressure during all examinations, whereas two GPs and one NP did only occasionally
Insert finger with pad posteriorly (task 26)	1.2 (1.3)	1.9 (1.9)	1.3 (.9)	-	0.7 (.7)	URs rarely executed this task whereas GPs and NPs spent longer than CRs
Assessment of sphincter tone (task 27)	19 (3.9)	-	-	-	-	Only performed by two NPs
Insert finger beyond sphincter into rectum (task 28)	2 (1.8)	2.7 (2.7)	2.5 (1.4)	1.3 (1)	1.8 (1.4)	GPs and NPs spend nearly twice as long as URs whereas CRs were between these groups.
	Across participants, they spent longer on N (2.7s) compared to UB (1.8s), BB (2s), UC (1.8s) and BC (1.9). The differences between N and UB (**p=.31**), UC (**p=.28**), and BC (**p=.008**) was significant; **F(4,67.9)=3.54** (Bonferroni-adjusted LMM).					
Coccyx palpation (task 29)	3.3 (4)	4.3 (5.4)	4.0 (6.4)	0.8	1.5 (1.8)	GPs and NPs took the longest whereas CRs and URs the shortest.
Rectal wall palpation (task 33)	6.4 (7.5)	5.8 (6.3)	10.3 (9.9)	1.2 (.9)	6.1 (4.1)	NPs spent twice as long as other cohorts: CRs, GPs and URs.
Prostate palpation (task 34)	20 (12.4)	23.5 (12.4)	18.7 (9.8)	18.3 (9.1)	20 (16.7)	GPs took slightly longer compared to CRs and NPs whereas URs took the shortest.
	Across participants, they spent longer on prostates difficult to identify: UB (23.1s), N (21.3s) and BB (20s) whereas shorter on other types including UC (19.1s) and BC (17.2s). The difference in duration between UB and BC was *significant*; **F(4,66.64)=3.2, p=.033** (Bonferroni-adjusted LMM). Total variability can be explained partly due to differences in clinicians (55.6%; **F(19,171)=12.5; p<.05**) and differences in prostate types (4.3%; **F(9,171)=2.05; p<.05**); **ICC(2,1)=.024 (.329)**.					
Internal examination of DRE	31.3 (20.2)	36.9 (24.1)	38.7 (20.3)	21.5 (10.2)	29.0 (19.6)	GPs and NPs took the longest overall, followed by CRs whereas URs spent the shortest.
	Across participants, the difference in duration between UB&BC significant; **F(4,65.55)=3.97, p=.029** (Bonferroni-adjusted LMM).					

Means (standard deviation) of duration are shown for each task and grouped by specialty. Findings from the sequence of tasks is summarised and indicated when relevant. *Clinical disciplines*: General Practitioner (GP), Nurse Practitioner (NP), Urologist (UR) and Colorectal surgeon (CR). *Types of prostate*: Normal (N), Unilateral Benign (UB), Bilateral Benign (BB), Unilateral Carcinoma (UC) and Bilateral Carcinoma (BC). Statistical differences are reported using a Linear Mixed Model (LMM).



Task	23	24	26	27	28	29	32	33	34	45
23		.16	.52		.32					
24			.87		.13					
26				.03	.97					
27					1.0					
28						.36		.17	.46	
29							.89			.11
32								.87	.13	
33						.11			.71	.18
34								.49		.51
45										

Fig. 8. Markov model (above) with tasks as states (start state in grey, end state in black, transitory states in white, encircled numbers refer to task numbers in Table I) and transition probabilities (below) computed from the average of number of tasks executed after another task across all participants.

TABLE V
SEQUENCE OF TASKS

Initial	Next	GPs	NPs	URs	CRs
28	34	.4(.5)	.26(.4)	.94(.2)	.26(.4)
33	34	.88(.7)	1.7(1.2)	.08(.3)	.94(1)
34	33	.44(.6)	1.56(1.1)	.2(.4)	.58(.9)
33	45	.1(.3)	.54(.5)	.02(.1)	.26(.4)
34	45	.9(.3)	.44(.5)	.82(.4)	.72(.5)

Means (standard deviation) of number of times a task is executed after another one across specialties.

pressure (Fig. 9 bottom). These frequencies, together with its power, are plotted in Fig. 10. for all participants' experiments.

The pressure applied per task by each participant is plotted in Fig. 11. For each participant, a pair of examinations of the same prostate type was considered consistent if it is similar within a 10%. A participant was not consistent if none or only one examination pair was consistent, consistent if two or three examination pairs were consistent and very consistent if four or five examination pairs were consistent.

Table VI shows a summary of the findings related to the pressure applied during each task with a level of stratification across specialties. From highest to lowest average standardized pressure, assessment of sphincter tone is the task during which participants applied the highest pressure (0.78), followed by prostate palpation (0.24), insert finger beyond sphincters into rectum (–0.22), coccyx palpation (–0.24), rectal wall palpation (–0.42), insert finger with pad posteriorly (–0.83) and apply gentle pressure on anus (–1.05).

Finally, a Spearman's rho two-tailed test was used to correlate the mean and standard deviation of the pressure applied on

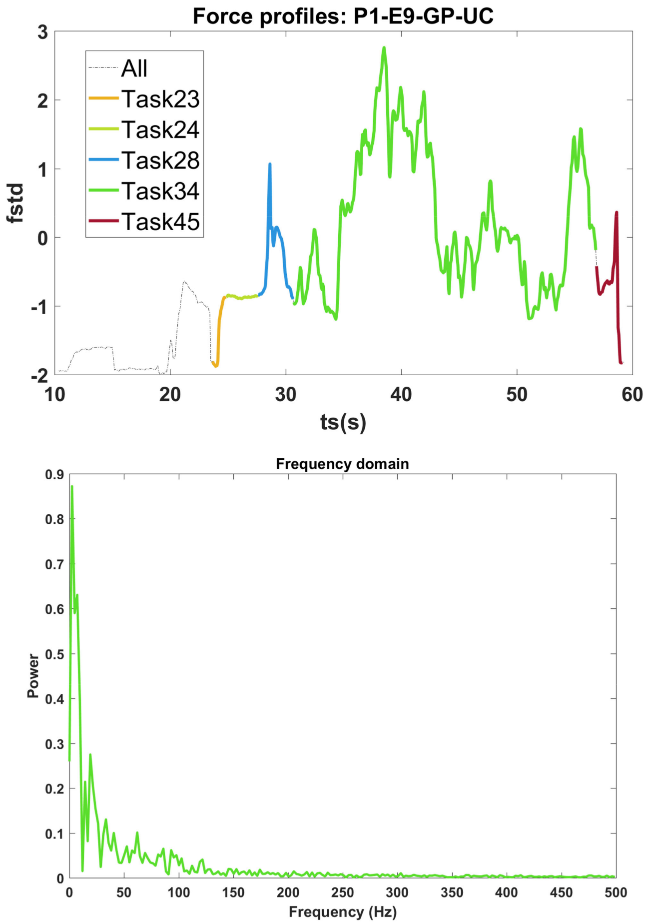


Fig. 9. **Top**: Task-annotated normalised force profiles (standard score) of an examination starting with finger on anus (orange), followed by applying gentle pressure (green), inserting finger with pad posteriorly (before peak in blue) and insert finger beyond sphincters into rectum (after peak in blue), prostate palpation (bright green) and finger removal (dark red). The standard score indicates the number of standard deviations away from the mean. This example is of an examination done by a GP on a unilateral carcinoma (UC) prostate type. **Bottom**: Frequency domain of pressure applied during prostate palpation (task 34).

the prostate during task 34 with correct diagnosis. No statistical significance was observed between mean pressure and correct diagnosis (rho = 0.092; p = 0.36; N = 100) nor between standard deviation of the pressure applied and correct diagnosis (rho = −0.166; p = 0.099; N = 100).

IV. Discussion

A. Diagnosis and Questionnaire

A physical examination such as DRE is a subjective process and in clinical practice its diagnosis relies on a full clinical history. Whilst DRE is much more difficult to perform on a patient than on a model, the results of the diagnosis are not surprising since established cancers are easier to diagnose, whereas small degrees of benign enlargement are easy to miss. Given that the benchtop model has large and obvious cancers, it is still surprising that five participants did not recognise it as such. In reality, cancers at diagnosis are likely to be smaller and more difficult to diagnose via palpation. Nonetheless, compared to previous studies of trainees doing DRE on three different simulators [14] with a percentage of correct identifications reported around 60%, we observed on average a lower sensitivity in our study with experts ($\mu = 54\%$, $\sigma = 24.37$), mostly due to misdiagnosed benign prostate types. This could be caused by participants with vast amount of expertise doing an examination on a fairly limited part-task trainer benchtop model (typically used by novices), by asking clinicians to emphasise the tasks that they commonly perform in practice and possibly by wearing the sensors used in our study. Additionally, 13.75% of unreported abnormalities (11/80, i.e., off-diagonal values of first column in Table II) appears to be consistent and slightly lower than that of previous studies with trainees (around 18%) [14].

Whilst GPs may not perform DREs as routinely as the other clinical specialities, DREs performed in a primary care setting are an important means of screening for serious prostate or anorectal abnormalities. They can be crucial in deciding if referral to a specialist for further studies is necessary. URs and CRs use the examination to inform clinical decision about what to do next, whether to perform follow-on examinations, or indeed an operation. NPs, who are typically urology/colorectal specialists, have a similar role as GPs in secondary care in conferring their judgements with URs and CRs. These rule-based decision making skills are particularly highlighted in previous work [14] during training of DRE in contrast to only physical palpation skills-based training.

Regarding the questionnaires, the lack of agreement as to whether the pressure sensor was comfortable to wear and did not affect participant's performance may be explained by the fact that, although thin, the pressure pad can still get in the way. Wearing the sensors and doing an examination on a benchtop model, which was reported to be very stiff, might explain why the examination was not completely comfortable.

B. Task Analysis

1) Duration: It is essential to put DRE in context to make sense of our findings. Whilst performing the routine parts of the examination, clinicians try to identify abnormalities, as well as reassure the patient and carefully decide how best to communicate their findings, taking into account their level of confidence in any diagnosis. Therefore, examination time is not only for palpation, but also for interpretation on the implications of palpation. Use of a plastic model precludes trainees from practicing this crucial aspect of a DRE.

Our findings related to duration ($\mu = 31.3$ s) across participants and regardless of their specialty, are similar to those reported in previous work [12], [14] of trainees and physicians, who typically utilised 30 seconds. Our results, however, highlight the differences of average duration, and their variability, with respect to type of specialty and task executed. For instance, both GPs and NPs spent longer and performed more tasks compared to URs and CRs. GPs are generalists who are likely to be the first to examine a patient presenting symptoms that dictate a DRE to be performed, whereas NPs tend to be more systematic and thorough since that is the approach they learn during training. URs and CRs routinely perform DREs on patients that

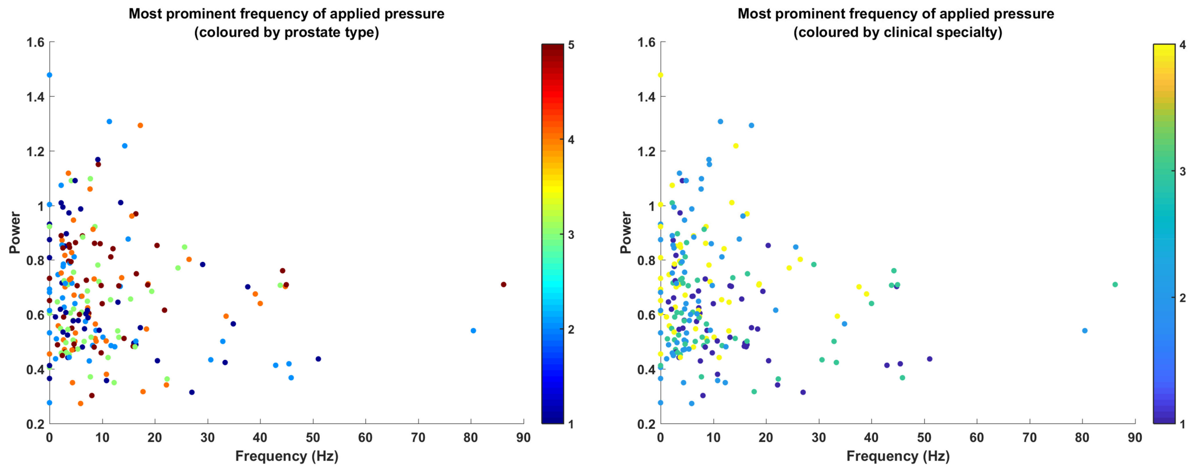


Fig. 10. Scatter plot of most prominent frequency (with its power) of applied pressure of all experiments coloured by type of prostate (left) and by clinical specialty (right). *Types of prostate*: Normal (1), Unilateral Benign (2), Bilateral Benign (3), Unilateral Carcinoma (4) and Bilateral Carcinoma (5). *Clinical disciplines*: General Practitioner (1), Nurse Practitioner (2), Urologist (3) and Colorectal surgeon (4).

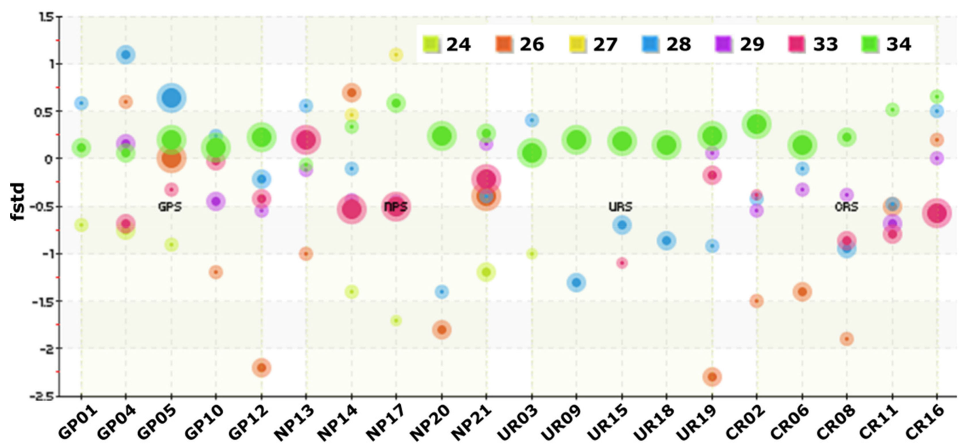


Fig. 11. Average normalised pressure by task (Table I) for all participants, grouped by speciality. Circle size indicates intra-subject consistency: *not consistent* (small), *consistent* (medium) and *very consistent* (large). Clinical disciplines: General Practitioner (GP), Nurse Practitioner (NP), Urologist (UR) and Colorectal surgeon (CR).

TABLE VI
SUMMARY OF FINDINGS (PRESSURE)

Task	Pressure					Notes
	Mean	GP	NP	UR	CR	
Apply gentle pressure (24)	-1.05 (.7)	-0.75 (.8)	-1.36 (.6)	-	-	
Insert finger with pad posteriorly (26)	-.83 (1.1)	-0.72 (1.2)	-0.55 (1.1)	-	-1.1 (.9)	URs inserted the finger with pad anteriorly (-2.27)
Assessment of sphincter tone (27)	.78 (.3)	-	-	-	-	
Insert finger beyond sphincter into rectum (28)	-.22 (.8)	0.46 (.7)	-0.36 (.8)	-0.66 (.7)	-0.29 (.6)	GPs applied the highest pressure followed by CRs and NPs whereas URs applied the lowest. Only the difference between GPs and URs was *significant*; **F(3,10.48)=5.1, p=.32** (Bonferroni-adjusted LMM).
Coccyx palpation (29)	-.24 (.6)	-0.22 (.7)	-0.29 (.6)	0.06	-0.47 (.5)	URs applied the highest pressure followed by GPs and NPs whereas CRs applied the lowest.
Rectal wall palpation (33)	-.42 (.5)	-0.37 (.4)	-0.39 (.4)	-0.21 (.5)	-0.68 (.4)	URs applied the highest pressure followed by GPs and NPs whereas CRs applied the lowest, although only one UR palpated the rectum.
Prostate palpation (34)	.24 (.3)	0.15 (.1)	0.27 (.4)	0.16 (.1)	0.38 (.3)	CRs applied the highest pressure followed by NPs whereas URs and GPs applied the lowest. NPs is the group varying the most, whereas URs the least.
	Across participants, variability of pressure applied resulted mostly from differences in clinicians (42.4%; **F(19,171)=7.7; p<.05**) and partly from differences in prostate types (8%; **F(9,171)=3.07; p<.05**); **ICC(2,1)=.058(.55).**					

Means (standard deviation) of pressure applied for all tasks from quantitative analysis of DRE. Pressure is shown as normalised based on the standard score. *Clinical disciplines*: General Practitioner (GP), Nurse Practitioner (NP), Urologist (UR) and Colorectal surgeon (CR).

have been referred and might therefore be expected to take less time performing it. Moreover, GPs/NPs would not normally have access to other tests/equipment that are commonly used by URs/CRs to inform their diagnosis.

Participants rarely applied gentle pressure before finger insertion. This may be explained by participants considering there was no need since it is a plastic model, the number of examinations they were asked to perform, or even due to a lack of clinical context. Regarding sphincter tone, clinicians may not routinely assess it, unless there is an indication for its assessment. Duration of prostate palpation was consistent with clinical practice, i.e., clinicians are likely to spend more time to confirm that no abnormality is present. Once an abnormality is detected, the examination might be shortened. GPs in particular may conduct further exploration even after abnormality detection, but at that point they already have enough information for a referral. This is confirmed by previous studies: "participants spend less time on a simulated examination when there is an obvious abnormality" [10].

2) Sequence: Overall, GPs/NPs were found to execute more steps. They would likely palpate the rectum after examining the prostate and may palpate the prostate again. Our findings suggest that URs rarely insert the finger posteriorly, seldom palpate the rectal walls, and concentrate mostly on palpating the prostate. However, in clinical practice (i.e., not on a benchtop model), this is likely to depend on the indications and clinical history of the patient. The coccyx was usually palpated after inserting the finger beyond the sphincters into the rectum, and then used as a landmark before starting a 360° sweep of the rectal walls.

The execution of tasks across clinical specialities can be summarised using a Markov model and transition states, which indicate the probability of executing a task based on the previous task. This state model can be used for comparing the performance of a novice against that of experts as a whole, or for a particular specialty. Also, the transition probabilities may be adapted to reflect different clinical scenarios (with a particular indication and clinical history), and therefore could be used to compare tasks performed by novices with tasks that are deemed to be essential to execute for better diagnosis.

C. Pressure

The normalisation of pressure data using the standard score was necessary to account for the variation in the calibration process and allow direct comparison across participants and specialities.

Sphincter tone assessment was the task with the highest pressure applied, although it was only performed by two participants (NPs). Our results also suggest that GPs applied more pressure compared to URs when their finger was inserted beyond sphincters into the rectum, a difference which was significant. During rectal wall palpation, CRs applied the least pressure compared to all other specialities. Related to prostate palpation, URs is the most consistent group and, together with GPs, applied lesser pressure compared to other specialities. Similar to the differences observed in pressure applied, inconsistent finger palpation techniques have been observed in previous studies [12], both within subjects and across different types of abnormalities. The fact that greater pressure was applied on the prostate compared to the rectum is confirmed by previous studies: "prostate palpation uses slightly more pressure than rectal wall palpation" [9]. Clinicians apply more pressure on the prostate to feel for consistency and make a judgement, whereas rectal palpation is more about identifying any possible abnormalities and therefore does not require applying considerable pressure.

As per previous studies ($rho = -0.099$; $p = 0.588$; $N = 32$) [14], we found no correlation between pressure applied and correct diagnosis. Average normalised applied pressure was used similarly to the Average Intentional Finger Pressure (AIFP) presented in [12] to see whether clinical specialty type (independent variable) was a significant predictor of applied pressure across the five types of prostate (dependent variables) during prostate palpation (task 34). We conducted a multivariate analysis of variance (MANOVA) test and found no significant difference (Wilks' Lambda with $p = 0.28$), i.e., there was no effect of specialty type on applied pressure. This is in contrast to [12], where they found an effect when comparing resident physicians with nurse practitioner students, indicating the differences in expertise.

The most prominent frequencies observed in our study are consistent with Dominant Frequency (DIFF) values of around 6 Hz reported in previous studies [11]. However, further work is necessary to correlate these findings with correct diagnosis and understand the causes of higher prominent frequencies, such as those found above 30 Hz in our study.

D. Limitations

Although this paper is limited by the small number of recruited experts (five per specialty) performing DREs on a plastic benchtop model (which is reported to have important limitations), our results build on previous qualitative (mainly cognitive tasks analysis) and quantitative work with the aim of understanding palpation skills amongst relevant clinical specialties. Further studies with a larger number of participants performing DRE either on improved benchtop models or on human subjects are possible with the techniques proposed here, with the possibility of studies on healthy volunteers or patients being an important contribution of our work.

V. Conclusion and Future Work

We presented a real-time visualisation and analysis system using position and pressure sensors located on the examining finger. An experimental study of 20 experts from four clinical disciplines allowed us to characterise palpation skills when conducting DREs on a benchtop model, in order to compare their performance, and to study execution of tasks and pressure applied.

Training of DRE requires more suitable tools for learning than those of existing benchtop models that provide no feedback or assessment of performance. Such learning tools may be underpinned by the body of knowledge obtained through studies such as that presented in this paper. For instance, the amount of pressure to apply is not currently incorporated into present training and novices have no way of knowing how much pressure they ought to apply.

GPs and NPs were found to execute more tasks and spend longer compared to URs and CRs. URs executed the least number of tasks and were the most consistent compared to other groups. NPs alternate palpation between prostate and rectum more often than other groups. After sphincter tone assessment, prostate palpation is the task with the greatest pressure, whereas rectal palpation uses the lowest pressure. The relative pressure applied during prostate palpation was similar across participants and very consistent amongst URs. The Markov model summarises task execution and could be used to compare performance of novices against that of experts. By adapting probability transitions based on different clinical scenarios, the model could be used to assess trainees on the essential tasks that need to be executed.

We plan to use logistic regression to study which metrics contribute to adequate performance. Our system can be extended to real subjects and other physical examinations and could be used as a learning tool by allowing novices to wear the sensors and receive real-time feedback.

Acknowledgment

The authors would like to thank the study participants and also thank the National Institute for Health Research Imperial Biomedical Research Centre for the Infrastructure support.

References

[1] A. Aubin *et al.*, "The seven-step palpation method: A proposal to improve palpation skills," *Int. J. Osteopathic Med.*, vol. 17, pp. 66–72, 2014.

[2] S. Ullrich and T. Kuhlen, "Haptic palpation for medical simulation in virtual environments," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 4, pp. 617–625, Apr. 2012.

[3] O. Epstein *et al.*, *Clinical Examination*, 4th ed. Amsterdam, The Netherlands: Elsevier, 2008.

[4] S. J. Lederman and R. L. Klatzky, "Haptic perception: A tutorial," *Attention, Perception Psychophys.*, vol. 71, no. 7, pp. 1439–1459, 2009.

[5] J. N. Howell *et al.*, "The virtual haptic back: A simulation for training in palpatory diagnosis," *BMC Med. Edu.*, vol. 8, no. 14, pp. 1–8, 2008.

[6] A. E. Bharucha and S. S. C. Rao, "An update on anorectal disorders for gastroenterologists," *Gastroenterology*, no. 146, pp. 37–45, 2014.

[7] R. K. Wong *et al.*, "The digital rectal examination: A multicenter survey of physician's and students' perceptions and practice patterns," *Amer. J. Gastroenterol.*, vol. 107, pp. 1157–1163, 2012.

[8] American Cancer Society, Atlanta, GA, USA, Prostate cancer: Early detection, 2013.

[9] N. Low-Beer *et al.*, "Hidden practice revealed: Using task analysis and novel simulator design to evaluate the teaching of digital rectal examination," *Amer. J. Surg.*, vol. 201, no. 1, pp. 46–53, 2011.

[10] R. Balkissoon *et al.*, "Lost in translation: Unfolding medical students' misconceptions of how to perform a clinical digital rectal examination," *Amer. J. Surg.*, vol. 197, no. 4, pp. 525–532, 2009.

[11] N. Wang *et al.*, "Using a prostate exam simulator to decipher palpation techniques that facilitate the detection of abnormalities near clinical limits," *Simul. Healthcare, J. Soc. Simul. Healthcare*, vol. 5, no. 3, pp. 152–160, 2010.

[12] N. Wang *et al.*, "Quantifying palpation techniques in relation to performance in a clinical prostate exam," *IEEE Trans. Inf. Technol. Biomed.,* vol. 14, no. 4, pp. 1088–1097, Jul. 2010.

[13] A. Granados *et al.*, "Real-time visualisation and analysis of internal examinations – Seeing the unseen," *MICCAI*, vol. 8673, pp. 617–625, 2014.

[14] G. J. Gerling *et al.*, "The design and evaluation of a computerized and physical simulator for training clinical prostate exams," *IEEE Trans. Syst., Man, Cybern. A, Syst. Human*, vol. 39, no. 2, pp. 388–403, Mar. 2009.