

Measurement and Stochastic Modeling of Kidney Puncture Forces

D. J. VAN GERWEN, J. DANKELMAN, and J. J. VAN DEN DOBBELSTEEN

Department of BioMechanical Engineering, Delft University of Technology, Mekelweg 2, 2628CD Delft, The Netherlands

(Received 7 August 2013; accepted 4 October 2013; published online 16 October 2013)

Associate Editor Joel D. Stitzel oversaw the review of this article.

Abstract—The development of needle insertion robots and training simulators requires knowledge of the forces that arise when a needle is inserted into soft-tissue. The present study aims to construct stochastic models of the force required to puncture a kidney using a trocar needle, based on measurements. To this end, a total of sixty insertions were performed into porcine kidneys (*ex vivo*), at constant speed, using a linear motion stage. Axial force was measured at the needle hub and an ultrasound probe moved with the needle to enable identification of anatomical structures. Two force peaks were observed for each tissue layer punctured, one caused by the tip and one by the edge of the cannula. Based on ultrasound data these double-peaks were classified into four groups, related to kidney capsule and internal structures. Group size varied from 7 to 55 double-peaks. Force peaks in each group were evaluated in terms of peak force and drop in force for both tip and cannula, and stochastic models were constructed that describe the multivariate distribution of these metrics. Peak forces in the capsule and internal structures ranged up to 2 N and 8 N, respectively. The resulting models can be used to simulate kidney puncture events in a variety of applications.

Keywords—Needle, Insertion, Puncture, Experiment, Ultrasound, Peak identification, Principal component.

INTRODUCTION

Theoretical models of needle-tissue interaction mechanics play a crucial role in the development of needle technology ranging from conventional needles to needle insertion robots. These models are usually deterministic in nature, although the needle-tissue interaction process typically exhibits considerable variability. This paper presents stochastic models, based on measured force data, that describe the variability in forces encountered during needle insertion

into soft-tissue. Although these models are widely applicable, our focus is on the application in a training simulator for needle interventions in the kidney.

Background

Percutaneous nephrostomy (kidney catheterization) is a complex clinical intervention in which a needle has to be inserted precisely into a specific region of the kidney. Like most clinical skills, the ability to perform this difficult procedure can only be acquired by practicing on real patients. This approach puts at risk both patient and clinician.

Simulator training could reduce this risk,^{32,33} allowing novices to tackle the first part of the learning curve safely before going into clinical practice. An important training challenge is to familiarize the novice with the forces required to advance the needle, and specifically with the high levels of variability that are typical for real patients. We believe this challenge can be met using a computer-based simulator with force-feedback based on stochastic models of tissue failure strength, i.e., models that take into account random variability.

The current study is aimed at constructing these stochastic models on the basis of forces measured during needle insertion.

Problem Description

The problem at hand is that we wish to model the axial forces that arise as the needle travels through the kidney as depicted in Fig. 1. These axial forces can be characterized by a variety of metrics, so, to understand which metrics are most useful, we need to consider the clinical procedure.

The goal in percutaneous nephrostomy is to place a catheter inside the renal pelvis so as to drain urine from the kidney. This catheter is placed with the help

Address correspondence to D. J. van Gerwen, Department of BioMechanical Engineering, Delft University of Technology, Mekelweg 2, 2628CD Delft, The Netherlands. Electronic mail: d.j.vangerwen@tudelft.nl

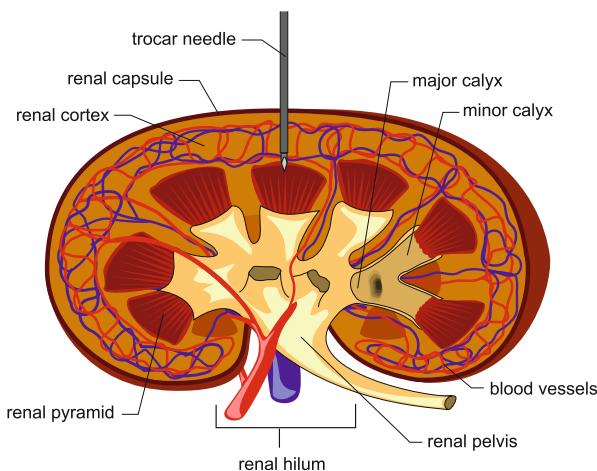


FIGURE 1. Schematic representation of kidney anatomy, showing the needle inserted towards renal pyramid (adaptation of illustration by Piotr Michal Jaworski, Creative Commons license CC BY-SA 3.0).

of a needle which is inserted into one of the calyces. The preferred approach is through the top-end of a calyx, as shown in Fig. 1, because this reduces the risk of blood vessel puncture.^{22,27} The needle therefore has to puncture the renal capsule and travel through the cortex and through a renal pyramid (medulla), possibly also traversing a calyx wall or blood-vessel wall.

The renal capsule consists mostly of connective tissue, and the calyces and pelvis consist of a combination of connective tissue and smooth muscle fibers, making them relatively tough compared with the cortex and medulla.^{12,26} When punctured, these tougher structures tend to produce a tactile “popping” sensation. It is common for clinicians to count the number of these sensations during the procedure, expecting one to occur as the needle passes the capsule and another when the needle enters a calyx.¹¹ The forces associated with these popping sensations are the ones that need to be characterized here.

Based on observation, the popping sensation is associated with an increase in force as the tissue is loaded, followed by a sudden decrease in force as the tissue fails. Thus we assume that a pop is characterized by a local force peak F , indicative of tissue failure strength, and a corresponding force drop ΔF , indicative of the strength of the popping sensation. The challenge is to model these two metrics for different tissue structures in the kidney, on the basis of force measurements.

Related Work

A survey by the authors³¹ showed that force data obtained from needle insertions into kidney tissue are scarce. Maurin *et al.*²⁴ inserted 18G trocar needles

(type not specified) into porcine kidney *in vivo*, both manually ($n = 5$) and automatically at 15 mm/s ($n = 4$). Healey *et al.*¹³ present a single force-position curve for automated insertion into porcine kidney *in vitro* at 8.3 mm/s (needle type and size not specified). Vidal *et al.*³² performed insertions up to 50 mm depth into porcine kidney, at 8.3 mm/s, using a Chiba needle (size unspecified). In addition, Zhai *et al.*³⁵ present data from a single insertion into porcine kidney, at 3 mm/s and up to 20 mm depth, using an 18 G trocar needle. None of these studies relate detailed force characteristics to specific parts of the kidney, nor do they evaluate the desired metrics.

The importance of modeling the stochastic nature of puncture forces was recognized by Kobayashi *et al.*¹⁹ who investigated the distribution of the puncture force metric F for unspecified 17 G needles inserted at speeds up to 8 mm/s into porcine liver (not kidney). They used a gamma distribution to model the puncture force, and found that the variability of the puncture force decreased with higher insertion speed.

Research Objective and Approach

The goal of the current investigation is to construct stochastic models of the peak axial force (F) and corresponding force drop (ΔF) resulting from the puncture of kidney capsule, calyx walls and pelvis, using a needle typical for the percutaneous nephrostomy procedure. To simplify implementation we base these models on well known parametric distribution functions, so the problem is reduced to one of parameter estimation. To reach this goal, we inserted trocar needles into porcine kidneys as discussed in detail below.

MATERIALS AND METHODS

Test Specimens

Porcine kidneys were used in the experiment, for practical reasons. Porcine kidneys are slightly larger than human kidneys, but are otherwise very similar in structure and function.^{25,30}

The porcine kidneys were obtained from the slaughterhouse the day before the experiment. They were excised immediately after death, and perinephric fat was removed, leaving the kidney capsule intact. The pigs were female Dutch Landrace Hybrids, ages approx. 4–6 months, weight approx. 90–110 kg. No distinction was made between left and right kidneys.

Within 8 h after death the kidneys were cast into blocks of high gel-strength Agar (Sigma-Aldrich, Zwijndrecht, the Netherlands), 2% concentration,

roughly $20 \times 10 \times 7$ cm in size, as depicted in Fig. 2. Gel temperature (before casting) did not exceed 45 °C. The specimens were stored at 4 °C, and before use the specimens were restored to room temperature (20 °C).

Measurement Protocol and Experimental Conditions

Three porcine kidneys were available in total. Each kidney was punctured in five random locations before switching to another kidney in random order. This was repeated up to a total of 60 insertions. All measurements were performed on the same day.

The kidneys, encased in rectangular blocks of gel, were placed on a rigid support submerged in water, and were oriented with the renal hilum pointed downwards, as depicted in Fig. 1. Needle orientation was fixed at 90° (i.e., vertical), but the insertion angle varied by as much as 30° depending on the kidney curvature at the insertion location. A fixed insertion speed of 10 mm/s was deemed appropriate on the basis of literature.³¹ The insertion distance was such that the needle passed completely through the kidney. This ensured that the capsule was punctured both proximally and distally during each insertion. Water temperature was maintained at 20 °C (room temperature).

Instrumentation

The measurement apparatus is depicted in Fig. 2, and a schematic overview of the experimental set-up can be found in the online supplementary materials accompanying this article. The set-up comprises a linear motion stage, mounted vertically on a frame composed of steel tubing, and equipped with a force/torque sensor and an ultrasound probe. The test specimens are placed directly below the stage and are submerged in water, together with the ultrasound probe, to ensure proper ultrasound transmission. The ultrasound probe is attached to the moving platform (similar to Dehghan and Salcudean⁹), so that it moves along with the needle during insertion.

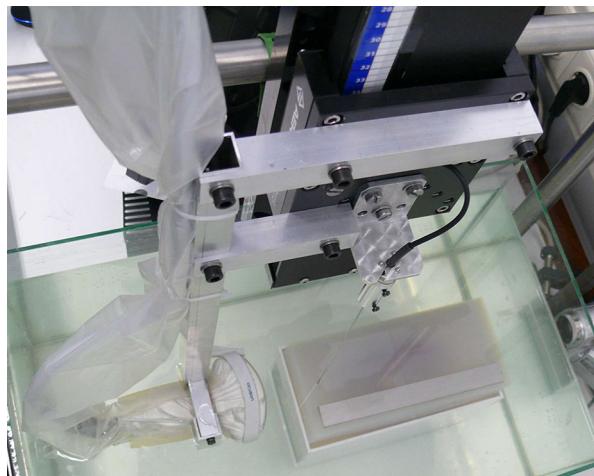


FIGURE 2. Linear motion stage with force sensor, needle, ultrasound probe, and test specimen encased in agar gel and submerged in water. Note that the ultrasound probe is in the out-of-plane configuration in this image, whereas during final measurements it was in the in-plane configuration.

torque sensor and an ultrasound probe. The test specimens are placed directly below the stage and are submerged in water, together with the ultrasound probe, to ensure proper ultrasound transmission. The ultrasound probe is attached to the moving platform (similar to Dehghan and Salcudean⁹), so that it moves along with the needle during insertion.

The linear stage is an Aerotech PRO115-400 with 400 mm travel and 5 mm/rev ball screw (Aerotech Inc., Pittsburgh, USA). Positioning-accuracy is in the order of 0.01 mm.

The needle hub is attached to an ATI nano17 six-axis force/torque sensor (ATI Industrial Automation, Apex, USA), with a 17 N rated load and resolution of 0.003 N in insertion direction, which is mounted on the moving platform of the linear stage.

Data acquisition and position control of the linear stage is achieved using a dedicated PC equipped with a dSPACE DS1104 real-time controller board (dSPACE GmbH, Paderborn, Germany). Forces and moments along all axes are stored on the computer, together with time stamp, platform position, and operating state (run, pause, retract). Analog data signals are sampled at 2 kHz with 12 bit resolution.

In addition to the force measurements, ultrasound images are captured at 18 fps (frames per second) using a dedicated video capture system (Noldus Information Technology, Wageningen, The Netherlands). An audio signal generated on the dSpace board is recorded with the ultrasound images for data synchronization purposes (operating states were characterized by specific audio frequencies). Ultrasound imagery is provided by an Acuson Sequoia 512 with convex Acuson 6C2 6 MHz probe (Siemens AG, Erlangen, Germany). The probe is oriented for an in-plane approach of the needle.

A HASCO Z251/1 digital thermometer (Hasco Hasenclever GmbH, Ldenscheid, Germany) is used to monitor the temperature of the water containing the test specimen.

Needle Characteristics

All measurements are performed with 18G Echotip Trocar Needles of 20 cm length (Cook Medical, Bloomington, USA). This type of needle is typical for the nephrostomy procedure,¹¹ and consists of a blunt 1.3 mm diameter cannula, with a special grating for increased ultrasound reflection (Echotip), and a 1.0 mm diameter stylet with diamond shaped cutting tip. The angle between longitudinal axis and grinding plane is approximately 11°, and the radius at the point is approximately 0.01 mm.

The shape of the tip cross-section varies along the length of the tip as depicted in Fig. 3 (regions A–D). In

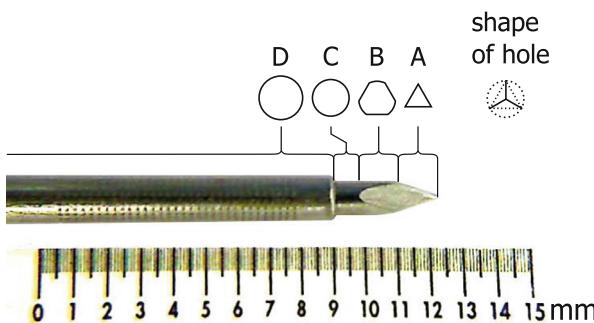


FIGURE 3. Close-up of the 18 G Echotip Trocar Needle with grating for increased ultrasound reflection. Regions with different cross-sections are indicated together with the shape of the resulting hole. Diameters: $c = 1.0\text{ mm}$, $d = 1.35\text{ mm}$.

a preliminary experiment we used synchronized video to investigate how this tip geometry influences the axial force characteristics during puncture of a thin polyurethane membrane. The result is depicted in Fig. 4.

Three distinct force peaks arise while puncturing this artificial membrane: one at the very tip (start of region A), where the needle starts cutting into the membrane, one near the transition from A to B, where the hole is wedged open, and one at the transition from C to D, where the membrane encounters the edge of the cannula. The second, smaller peak was not always present.

Based on this puncture force characteristic, the trocar needle is expected to produce at least two force peaks for every membranous tissue layer punctured: one due to the tip and one due to the cannula edge. We refer to this as a double-peak. To allow objective identification of these double-peaks, the force signal needs to be pre-processed.

Force Signal Processing

To prepare the measured force signal for objective identification of relevant peak forces, a 10-point double-pass moving average filter is applied (peak attenuation $<0.1\text{ N}$), followed by a top-down line-simplification algorithm based on the Douglas–Peucker algorithm.¹⁰

In essence, this algorithm reduces the number of data points used to represent a signal while retaining its most striking features (peaks and valleys). An error threshold value, ε , determines the relevant peak size.

The algorithm seeks to approximate the original signal using a subset of its data points. Initially the entire signal is approximated by one line segment, connecting the first and last point. The algorithm then starts adding data points to this list until a specified approximation accuracy is reached. In each iteration, the data point that has the greatest normal distance to the approximation line is added to the list. In our implementation, the dimensionless normal distance

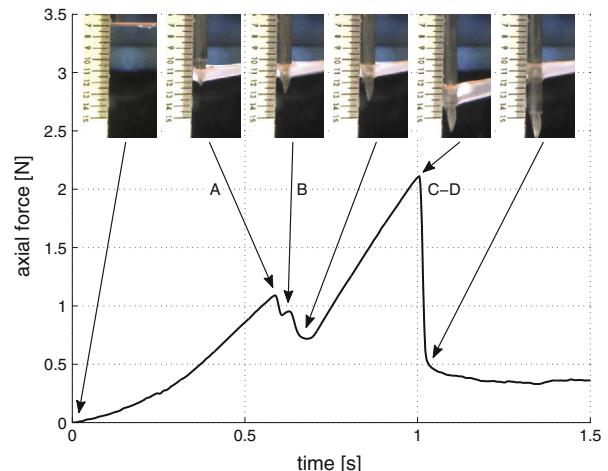


FIGURE 4. Typical force characteristic for the 18 G Cook diamond tip trocar needle (puncturing a 0.1 mm thick polyethylene membrane under constant uniaxial tension, at an insertion speed of 10 mm/s). A–D refer to the regions from Fig. 3. Similar characteristics for beveled needles are described in the DIN 13097-4 standard.

(l^2 -norm) between the line segments and the corresponding data points is used as the error criterion ε .

Force and position are normalized with respect to the force sensor rating (17 N) and the needle length (200 mm), respectively, so that ε is dimensionless. The diameter of the cutting part of the tip (1.0 mm) is assumed to be a useful characteristic length for the required level of detail. Thus, we let the error threshold equal the dimensionless tip diameter, i.e., $\varepsilon = \frac{1.0}{200} = 0.005$. For this threshold value, the number of points in the signal is typically reduced by three orders of magnitude (from 10^4 to 10) as in Fig. 6.

The force peaks that remain after simplification are cross-checked with the ultrasound data.

Ultrasound Signal Processing

Synchronization of ultrasound data and force data is achieved using the audio signal from the control PC, which was recorded with the ultrasound data. This is done automatically by detecting frequency changes in the audio signal with the help of fast Fourier transforms. Given the ultrasound frame rate of 18 fps , a synchronization uncertainty of approximately 0.06 s (i.e., 0.6 mm at 10 mm/s) has to be taken into account.

After synchronization, the ultrasound frames corresponding to the vertices of the simplified force signal are extracted for visual inspection.

Peak Classification

Membrane puncture events are identified by visual inspection of both the selected ultrasound images and

the simplified force-position data. These puncture events are expected to manifest in the force-position data as double-peaks similar to the one depicted in Fig. 4. Only distinct double-peaks are taken into account. Single force peaks or multiple peaks that are not clearly identifiable, i.e., not distinct and not similar to Fig. 4, are not considered in the analysis.

Double-peaks occur either when the needle moves into the capsule or some internal structure, or when it moves out of one of those structures. Thus, based on the ultrasound information, each double-peak is classified into one of four groups: (*ci*) capsule in, (*ii*) internal structure in, (*io*) internal structure out, or (*co*) capsule out.

Definition of Force Metrics

For each double-peak identified, four metrics are evaluated: the tip-peak F_t , tip-drop ΔF_t , cannula-peak F_c , and cannula-drop ΔF_c . These metrics are defined in Fig. 5.

The peak forces are defined with respect to an estimated friction baseline, which is represented by vertex 1 in the figure, i.e., the estimated point of initial contact between the needle tip and the membranous tissue layer.

Since F_t occurs first in time, it is treated as independent. The other metrics are then dependent by definition. For example, $F_t = F_2 - F_1$ and $F_c = F_4 - F_1$, so that $F_c = F_t + F_4 - F_2$. As a result, both ΔF_t and F_c have a component that is directly proportional to F_t , and ΔF_c has a component that is directly proportional to F_c .

There are also physical reasons to expect association between the metrics. For example, it is plausible that the quality of the cut made by the tip depends on F_t and that it will influence both F_c and ΔF_c , as well as the remaining friction after puncture (vertex 5). Similarly, tissue toughness is a probable confounding factor, as it

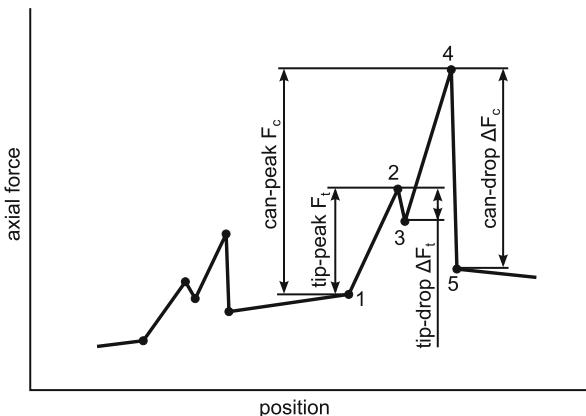


FIGURE 5. Definition of force metrics and vertices 1–5.

can be expected to influence both F_t and F_c , and may also influence the corresponding drops in force.

Sample correlations between metrics are quantified using Spearman's rank correlation r_s , which is more robust than Pearson's coefficient.²⁹

Stochastic Modeling of the Force Metrics

As stated in the introduction, our goal is to model the observed force metrics in terms of parametric distribution functions. For each tissue class *ci*, *co*, *ii*, and *io*, we wish to establish a multivariate model that describes the distribution of the four metrics F_t , ΔF_t , F_c , and ΔF_c .

The modeling approach adopted here is to apply a coordinate transformation that reduces the sample correlations to zero, and then model each new variable as an independent random variable (the disparity between correlation and dependence is discussed later). This approach involves three steps (a flow diagram is provided in the electronic supplementary materials):

Step 1 is to determine the orthogonal rotation matrix P (4×4) that transforms the original correlated variables $\mathbf{X} = [F_t \ \Delta F_t \ F_c \ \Delta F_c]^T$ into a new set of uncorrelated variables \mathbf{Y} , where $\mathbf{Y} = P\mathbf{X}$. This is achieved by principal component analysis (PCA).²¹ The columns of P are called the principal components (PC1 to PC4), and they consist of the unit eigenvectors of the sample covariance matrix.

Step 2 is to apply the coordinate transformation (rotation) $\mathbf{Y} = P\mathbf{X}$ to the observed data, thereby removing the sample correlations.

Step 3 is to select parametric distribution functions for the transformed data and estimate their parameters. This is done independently for each of the new variables in \mathbf{Y} , using either the normal distribution function $\mathcal{N}(\mu, \sigma)$ or the lognormal distribution function $\ln \mathcal{N}(\mu, \sigma)$. If the transformed observations take negative values, a normal distribution function is used. Otherwise, the choice between normal and lognormal distribution function depends on the best goodness-of-fit in terms of the Anderson-Darling statistic^{7,3} (a measure of discrepancy between theoretical and empirical distribution functions). Parameters μ and σ are estimated in the maximum likelihood sense.

The final model for each tissue class (*ci*, *co*, *ii*, or *io*) now consists of a transformation matrix P , and the distribution types and corresponding estimated parameters $\hat{\mu}$ and $\hat{\sigma}$ for each of the four transformed variables.

Simulation of the Force Metrics

To simulate puncture events using the proposed models, the process is reversed. This involves two steps:

Step 4 is to independently draw sets of (quasi-) random numbers from each of the estimated distributions (either $\ln \mathcal{N}(\hat{\mu}, \hat{\sigma})$ or $\mathcal{N}(\hat{\mu}, \hat{\sigma})$). This yields a simulated dataset in terms of the transformed variables \mathbf{Y} .

Step 5 is to transform this simulated dataset back to the original variables, using the inverse rotation $\mathbf{X} = P^T \mathbf{Y}$. Negative values are discarded, as discussed later.

To assess the correspondence between the proposed models and the original data, we conduct a Monte-Carlo study using $N = 10^5$ random datasets of size $4 \times n$ for each group. The two-sample Kolmogorov-Smirnov (KS) test^{15,23} is used, for each metric individually, to determine the probability that the simulated data and measured data are samples from the same probability distribution. A conservative significance level of $\alpha = 0.01$ is employed to account for the fact that four metrics are tested.

RESULTS

Raw Data

The absolute maximum force measured during the entire experiment was 12 N. An example of raw force-position data from a single kidney insertion is presented in Fig. 6, together with its simplified representation. This figure shows a number of double-peaks that are very similar in shape to the one depicted in Fig. 4.

The double peaks described in the “Materials and Methods” section are easily recognized by the human eye, but a set of rules by which to identify and classify these peaks automatically was difficult to establish. Therefore, classification was achieved by inspection of

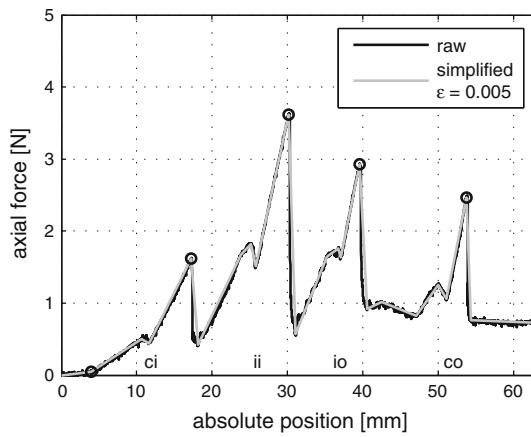


FIGURE 6. Example of raw force-position data and the corresponding simplified representation. Circles correspond to the ultrasound frames in Fig. 7. Peaks are classified as *ci* (capsule in), *ii* (internal in), *io* (internal out), *co* (capsule out), based on ultrasound data.

the simplified force response, as depicted in Fig. 6, in conjunction with ultrasound data.

Figure 7 shows the sequence of ultrasound frames corresponding to the markers in Fig. 6. The first frame represents the point of contact between needle and kidney, and subsequent frames show the point of maximum displacement for each membranous layer, i.e., the point where the cannula edge pops through (vertex 4 in Fig. 5).

In the ultrasound images, the renal capsule is recognized as a thin white (hyper-echoic) line enclosing the cortex. In the center we recognize some dark spots surrounded by white lines that are likely to belong to the collection system. Based on these images, it is clear that some kind of internal structure is pierced, but it is difficult to determine the specific type or part of that structure (e.g., calyx, pelvis, blood vessel). Thus we are limited to a crude distinction between capsule and internal structures.

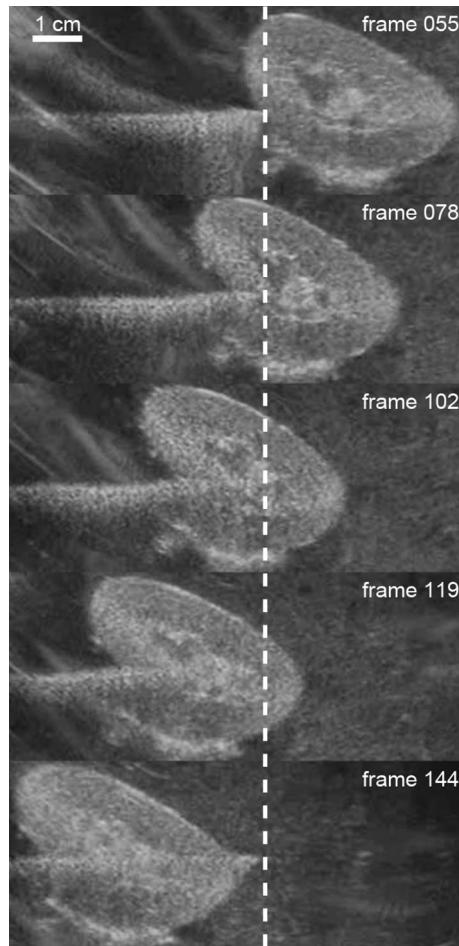


FIGURE 7. Sequence of ultrasound frames corresponding to the positions highlighted in Fig. 6. The dashed line represents the (horizontal) position of the needle tip, which is fixed with respect to the ultrasound probe.

Judging from the force data in Fig. 6, it seems reasonable to state that the needle punctures four distinct membranous structures. The ultrasound images confirm that these puncture events correspond to capsule entry (*ci*), entry into and exit from some internal structure (*ii* and *io*), and capsule exit (*co*), respectively.

However, only four measurements were as clear-cut as the current example. In eight cases, double-peaks appear only for the capsule, suggesting that no internal structures were encountered. In forty-eight cases, one or at most two double-peaks can be identified, whereas the other peaks are discarded because they cannot be distinguished from each other. Examples of the great variety of force-responses are provided in the electronic supplementary materials. The resulting sample sizes, i.e., the number of double-peaks identified, are $n_{ci} = 55$, $n_{co} = 47$, $n_{ii} = 10$, and $n_{io} = 7$. For each of these groups we evaluate the force metrics.

Force Metrics

Figure 8 shows scatter plots for all combinations of metrics for the *ci* and *co* groups (capsule), with corresponding sample correlations r_s . In both groups, ΔF_t vs. F_t and ΔF_c vs. F_c show considerable (linear) correlation, whereas the other combinations do not. In both groups we find that F_c is mostly larger than F_t , ΔF_t is considerably smaller than F_t , and ΔF_c is approximately equal to F_c . The *co* group shows greater variability than *ci*.

Similar scatter plots for the *ii* and *io* groups (internal structures) can be found in the electronic supplementary materials. All sample correlations for *ii* and *io* are large ($r_s > 0.68$), but none are statistically significant. Again, for most observations, $F_c > F_t$, $\Delta F_t < F_t$, and $\Delta F_c \approx F_c$.

Comparing the capsule groups (*ci* and *co*) with the internal-structures groups (*ii* and *io*), we find that ΔF_t takes similar values in both capsule and internal structures, but the other metrics reach much higher values in the internal structures (e.g., $F_{c,io} < 8$ N whereas $F_{c,co} < 2$ N).

Stochastic Models of the Force Metrics

The principal component matrices (P) and estimated distributions ($\mathcal{N}(\hat{\mu}, \hat{\sigma})$ or $\ln \mathcal{N}(\hat{\mu}, \hat{\sigma})$) for groups *ci* and *co* are presented in Tables 1 and 2. Those for *ii* and *io* are found in the electronic supplementary materials. The tables also show the percentage of total variance explained by each of the principal components. Note that the parameters for normal distributions and lognormal distributions cannot be directly compared, because the latter are dimensionless whereas the former are in Newtons.

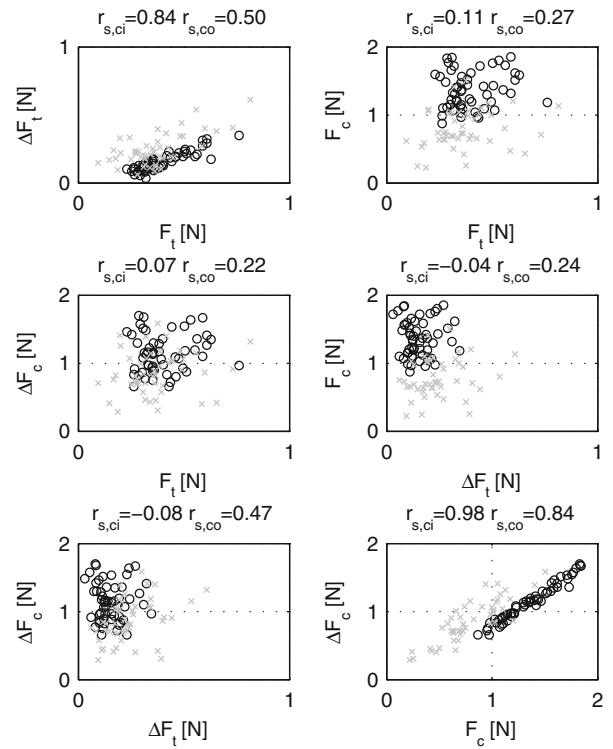


FIGURE 8. Scatter plots showing all possible combinations of metrics for the *ci* group (○) and *co* group (×), with corresponding Spearman rank correlations r_s ($n_{ci} = 55$ and $n_{co} = 47$).

Simulation Results

The Monte-Carlo performance evaluation ($N = 10^5$) showed that 1.6% of the simulated observations were rejected by the two-sample Kolmogorov-Smirnov test, at a significance level $\alpha = 0.01$, as not being likely to come from the same distribution as the original dataset. For illustrative purposes, a figure is included in the electronic supplementary materials that shows a representative example of a simulated dataset for *ci*, together with the original observations.

DISCUSSION

Summary of Main Results

Classification of force peaks proved challenging due to qualitative variability in force response and limited quality of ultrasound footage, and resulted in different group sizes: $n_{ci} = 55$, $n_{co} = 47$, $n_{ii} = 10$, and $n_{io} = 7$ (from a total of 60 insertions). Total axial force observed during the experiments did not exceed 12 N. For the capsule groups (*ci* and *co*), $F_t < 1$ N, $\Delta F_t < 0.5$ N, $F_c < 2$ N, and $\Delta F_c < 2$ N, whereas for the internal structure groups (*ii* and *io*) $F_t < 3$ N, $\Delta F_t < 0.5$ N, $F_c < 8$ N, and $\Delta F_c < 8$ N. For all groups it was found

TABLE 1. Descriptive model for the *ci* group (*n* = 55).

Variance explained	PC1 88.6%	PC2 10.2%	PC3 0.8%	PC4 0.4%
P_{ci}	0.043 0.001 0.691 0.722	0.863 0.498 0.030 −0.080	0.119 −0.350 0.668 −0.646	−0.489 0.793 0.275 −0.236
Distribution	$\ln \mathcal{N}$	$\ln \mathcal{N}$	$\ln \mathcal{N}$	\mathcal{N}
$\hat{\mu}_{ci}$	0.552	−1.046	−1.833	0.033
$\hat{\sigma}_{ci}$	0.222	0.343	0.224	0.026

TABLE 2. Descriptive model for the *co* group (*n* = 47).

Variance explained	PC1 79.1%	PC2 12.7%	PC3 6.9%	PC4 1.3%
P_{co}	0.088 0.105 0.660 0.739	0.817 0.557 −0.141 −0.050	0.323 −0.359 0.659 −0.576	−0.470 0.741 0.332 −0.346
Distribution	\mathcal{N}	$\ln \mathcal{N}$	\mathcal{N}	\mathcal{N}
$\hat{\mu}_{co}$	1.211	−1.434	0.059	−0.026
$\hat{\sigma}_{co}$	0.412	0.590	0.122	0.052

that $F_t \lesssim F_c$, $\Delta F_t \lesssim F_t$, and $\Delta F_c \approx F_c$. High sample correlations were observed between ΔF_t and F_t , and between ΔF_c and F_c in all groups.

The multivariate distribution of the metrics was modeled using principal component matrices combined with either normal or lognormal distribution functions. An assessment of correspondence between modeled and measured data, using Monte-Carlo simulations, showed a 1.6% rejection rate for the two-sample KS-test.

Interpretation of Main Results

The capsule-in group (*ci*) is easily recognized in most cases, even without ultrasound, and is therefore most reliable, with minimal error in terms of estimated friction baseline. Classification proved much more challenging for the remaining groups, as illustrated by the differences in group size. Most observations in the internal structures groups (*ii* and *io*) are likely to represent calyx wall punctures, but these observations should be treated with the proper caution.

The difference in magnitude of the force metrics between the capsule and internal structures suggests that the latter are generally harder to penetrate.

The metrics F_t and F_c can be interpreted as measures of tissue failure strength for two different load cases. The fact that $F_t \lesssim F_c$ is in line with expectation,

because the area of needle-tissue contact at the tip is much smaller than at the cannula edge, so that the failure stress is reached at a lower axial force.

A higher failure stress (stronger tissue) could be expected to increase both F_t and F_c , thus correlating the two. The high sample correlations for the internal structures groups seem to agree with this expectation, but these correlations were not statistically significant. On the other hand, no obvious correlation was observed in the (more reliable) capsule groups. It is possible that other factors, such as quality of the incision made by the tip, counteract the expected relation.

The association observed between ΔF_t and F_t is in agreement with the “**Definition of Force Metrics**” section and also with the nonlinearity of the force-position curve during membrane loading (e.g., Fig. 4). After initial membrane failure, represented by F_t , the membrane will unload until it encounters the edge of the cannula. The unloading distance (distance from tip to edge of the cannula) is constant. Due to the nonlinearity of the loading curve, the change in force (ΔF_t) associated with this distance will increase with F_t .

The association between ΔF_c and F_c also agrees with expectation: After final membrane failure, represented by F_c , the membrane will unload until its elastic energy is no longer sufficient to overcome friction. If the friction force is independent of F_c , it is reasonable to expect

ΔF_c to be directly proportional to F_c (in the extreme case of zero friction, ΔF_c would be equal to F_c).

The Monte-Carlo performance evaluation of the proposed models suggests that they provide an adequate description of the data. A rejection rate of 1.6% at $\alpha = 0.01$ suggests good correspondence between the individual distributions of the simulated metrics and the original metrics. Qualitative comparison of original and simulated scatter plots showed good general agreement.

Generalization of the Results

The question whether our findings translate to other settings, such as the clinical one, cannot be answered conclusively at this point. Nevertheless, some support is found in the literature and in previous experience.

The double-peak characteristic was also found in an unpublished pilot experiment in which we used the same type of needle to puncture the kidney and liver of a human cadaver *through the skin*, as illustrated by examples provided in the electronic supplementary materials. It is noted, however, that the tip-peak was often less pronounced in these measurements, especially at locations deep inside the body. Cannula-peaks in the human cadaverous kidney were of the same order of magnitude as those presently found in isolated porcine kidneys, but this observation should be treated with due caution: The comparison between isolated porcine and human kidneys should be addressed in a dedicated experiment.

In this respect, Snedeker *et al.*²⁸ compared *human* and *porcine* kidney capsule samples *ex vivo*, and found that mean capsule thickness (approx. 0.05 mm) and nominal ultimate stress in quasi-static uniaxial tension (in the order of 10 MPa) were similar in the two groups. They also found that both apparent stiffness and ultimate stress increased considerably with strain rate. Brunon *et al.*⁴ compared human and porcine *liver* capsule and found different mean thicknesses (approx. 0.1 mm and 0.2 mm respectively), but comparable ultimate stress (approx. 2 MPa) between the two. These uniaxial-tension test results are not directly comparable to the failure strengths observed during needle puncture, but they do suggest a reasonable correspondence in general mechanical properties between human and porcine kidney and liver capsules.

The double-peak characteristics are also expected to occur in the *in vivo* situation. The distribution of F_c in the *ci* (capsule-in) group of the current study (Fig. 8) is in good agreement with the results reported by Maurin *et al.*²⁴ who found a maximum puncture force of 1.22/0.34 N (mean/std, $n = 4$, 18 G needle of unknown type inserted at 15 mm/s) during *in vivo* porcine kidney puncture, under direct access.

No studies were found that investigate the influence of decay time on capsule *puncture* strength, but Hebert *et al.*¹⁴ compared the ultimate *tensile* strength of fresh canine kidney capsule samples with that of 30 h old samples, and found no significant effect. Stiffness, on the other hand, was found to increase within minutes after death in porcine heart samples,⁵ and Kerdok *et al.*¹⁸ found that general viscoelastic properties of porcine liver (stiffness and damping) increase noticeably in unperfused conditions, whereas those of perfused livers remain comparable to the *in vivo* situation. These findings could indicate that decay time is not critical for the puncture strengths of kidney capsule, but that the use of perfused organs could provide more realistic stiffness values and boundary conditions. In addition, it is expected that perfused conditions will improve ultrasound contrast, thereby enabling more reliable tissue identification and peak classification.

Limitations

The use of an ultrasound probe moving with the needle proved to be very useful in verifying contact with different tissue structures, because the location of the tip in the image is known even if the tip itself is not visible. However, the approach should be refined in order to allow more precise identification of internal structures.

Difficulties in classification caused the sample sizes for the internal structures groups to remain very small. This limits the reliability of the estimated models for *ii* and *io*.

Although the choice of threshold value (ϵ) for line simplification ensures reproducibility, it is inherently subjective, and will influence the number of force peaks taken into account in the analysis. However, peak identification inevitably requires such choices.

The algorithm used for line simplification was first proposed by Douglas and Peucker¹⁰ (specifically Method 2 from this reference). This method does not provide optimal simplifications^{1,20} (in terms of minimal error for a given number of points or vice versa), and, due to the top-down approach, it is intrinsically limited to off-line application.¹⁷ However, precisely because of the top-down approach, the method is known to produce good qualitative approximations.⁶ More recent developments in the field of line simplification focus mainly on reducing computational complexity, enabling on-line simplification, and ensuring optimality.^{1,2,16} Neither of these developments seems particularly important for the present study, because our analysis is performed off-line, and our primary goal is retention of salient points regardless of optimality. It is possible that optimal simplifications or different error measures will yield better retention of salient points, but this is expected to depend strongly on the dataset.

Concerning the stochastic modeling approach, it is noted that the use of principal components to remove correlation between variables does not cause them to become independent. However, the scatter plots of the observed data do not suggest obvious hidden dependencies.

The choice of normal and lognormal distribution functions was based on convenience. In contrast, Kobayashi *et al.*¹⁹ used a gamma distribution to model puncture forces, and in light of failure strength analysis a Weibull distribution might also be suitable.^{8,34}

The proposed modeling approach does not preclude negative values, although the force metrics are all strictly positive by definition. The effect of discarding negative values on the distribution shape is deemed negligible for the current dataset.

Insertion location was not taken into account in the present study, but there are indications that it may be an important factor. For example, Hebert *et al.*¹⁴ found large differences in local thickness, stiffness, and nominal ultimate stress in uniaxial tension, between different regions of individual canine kidney capsules.

CONCLUSION

The goal of this investigation was to construct stochastic models of the peak forces and subsequent drops in force that arise during puncture of tissue structures in the kidney. This goal was achieved using measurement data obtained under ultrasound observation with a moving probe. Reliable classification of peak forces proved a challenge despite the use of ultrasound, especially for internal kidney structures. Nevertheless, a useful dataset was obtained, and multivariate stochastic models were successfully constructed. These models can be used to simulate kidney puncture events in a variety of applications such as training, path planning, and robotics.

The basic methodology presented here might be extended to other problems involving isolated organs. A generic approach would involve the following steps: (1) Obtain a characteristic membrane puncture force response (e.g., Fig. 4); (2) Use this response to define objective metrics that are relevant for the specific problem (e.g., Fig. 5); (3) Select an appropriate line simplification tolerance depending on the dataset and metrics; (4) Classify the resulting observations based on ultrasound or other visualization methods; (5) Model the classified observations in the PCA reference system, using appropriate distribution functions. In our opinion steps 1 and 2 are particularly important.

ACKNOWLEDGMENTS

The authors would like to thank Sjoerd van Tuijl from LifeTec Group, and Kirsten Henken and Arjan van Dijke from Delft University of Technology for their assistance in the experiments.

CONFLICT OF INTEREST

There are no conflicts of interest to report.

REFERENCES

- ¹Abam, M. A., M. de Berg, P. Hachenberger, and A. Zarei. Streaming algorithms for line simplification. *Discret. Comput. Geom.* 43:497–515, 2010.
- ²Agarwal, P. K., S. Har-Peled, N. H. Mustafa, and Y. Wang. Near-linear time approximation algorithms for curve simplification. *Algorithmica* 42:203–219, 2005.
- ³Anderson, T. W., and D. A. Darling. A test of goodness of fit. *J Am. Stat. Assoc.* 49:765–769, 1954.
- ⁴Brunon, A., K. Bruyère-Garnier, and M. Coret. Mechanical characterization of liver capsule through uniaxial quasi-static tensile tests until failure. *J. Biomech.* 43:2221–2227, 2010.
- ⁵Choy, Y. B., H. Cao, S. Tungjikusolmun, J.-Z. Tsai, D. Haemmerich, V. R. Vorperian, and J. G. Webster. Mechanical compliance of the endocardium. *J. Biomech.* 35:1671–1676, 2002.
- ⁶Clayton, V. H. Cartographic generalization: A review of feature simplification and systematic point elimination algorithms. Tech. Rep. NOS 112, National Oceanic and Atmospheric Administration, 1985.
- ⁷D’Agostino, R. B., and M. A. Stephens. Goodness-of-Fit Techniques. New York: Marcel Dekker Inc., 1986.
- ⁸Dally, J. W. Statistical analysis of experimental data. In: Springer Handbook of Experimental Solid Mechanics, edited by W. N. Sharpe. New York: Springer, 2008, pp. 259–280.
- ⁹Dehghan, E., and S. E. Salcudean. Needle insertion study using ultrasound-based 2d motion tracking. In: Lect Notes Comput. Sci., vol. 5242, edited by D. Metaxas. Berlin: Springer-Verlag, 2008, pp. 660–667.
- ¹⁰Douglas, D. H. and T. K. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Can. Cartogr.* 10:112–122, 1973.
- ¹¹Dyer, R. B., J. D. Regan, P. V. Kavanagh, E. G. Khatod, M. Y. Chen, and R. J. Zagoria. Percutaneous nephrostomy with extensions of the technique: Step by step. *RadioGraphics* 22:503–525, 2002.
- ¹²Hall-Craggs, E. C. B. Anatomy—As a Basis for Clinical Medicine, 3rd edn. London: Williams & Wilkins, 1995.
- ¹³Healey, A. E., J. C. Evans, M. G. Murphy, S. Powell, T. V. How, D. Groves, F. Hatfield, B. M. Diaz, and D. A. Gould. In vivo force during arterial interventional radiology needle puncture procedures. In: St Heal T, vol. 13, edited by J. Westwood. IOS Press, 2005, pp. 178–184.
- ¹⁴Hebert, L. A., W.-C. Chen, A. Hartmann, and J. C. Garancis. Mechanical properties of the dog renal capsule. *J. Appl. Physiol.* 40:164–170, 1976.

- ¹⁵Hodges, J. L. The significance probability of the Smirnov two-sample test. *Arkiv för Matematik* 3:469–486, 1958.
- ¹⁶Imai, H., and M. Iri. Polygonal approximations of a curve-formulations and algorithms. In: Computational Morphology, edited by G. T. Toussaint. Amsterdam: Elsevier, 1988, pp. 71–86.
- ¹⁷Keogh, E., S. Chu, D. Hart, and M. Pazzani. Segmenting time series—a survey and novel approach. In: Data Mining in Time Series Databases, edited by M. Last, A. Kandel, and H. Bunke. Singapore: World Scientific Publishing, 2004, pp. 1–21.
- ¹⁸Kerdok, A. E., M. P. Ottensmeyer, and R. D. Howe. Effects of perfusion on the viscoelastic characteristics of liver. *J. Biomech.* 39:2221–2231, 2006.
- ¹⁹Kobayashi, Y., A. Onishi, T. Hoshi, and K. Kawamura. Modeling of conditions where a puncture occurs during needle insertion considering probability distribution. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France, 2008, pp. 1433–1440.
- ²⁰Kolesnikov, A., and P. Fränti. Reduced-search dynamic programming for approximation of polygonal curves. *Pattern Recogn. Lett.* 24:2243–2254, 2003.
- ²¹Lay, D. C. Linear Algebra and Its Applications, 2nd edn. Reading, MA: Addison-Wesley, 2000.
- ²²Marcovich, R., and A. D. Smith. Percutaneous renal access: tips and tricks. *Brit. J. Urol. Int.* 95:78–84, 2005.
- ²³Massey, F. J. The kolmogorov-smirnov test for goodness of fit. *J. Am. Stat. Assoc.* 46:68–78, 1951.
- ²⁴Maurin, B., L. Barbe, B. Bayle, P. Zanne, J. Gangloff, M. De Mathelin, A. Gangi, L. Soler, and A. Forgione. In vivo study of forces during needle insertions. In: Medical Robotics, Navigation and Visualisation Scientific Workshop, Remagen, Germany, 2004, pp. 415–422.
- ²⁵Pereira-Sampaio, M. A., L. A. Favorito, F. J. B. Sampaio. Pig kidney: Anatomical relationships between the intrarenal arteries and the kidney collecting system. Applied study for urological research and surgical training. *J. Urology* 172:2077–2081, 2004.
- ²⁶Schuenke, M., E. Schulte, and U. Schumacher. Thieme Atlas of Anatomy—Neck and Internal Organs, 1st edn. Stuttgart: Georg Thieme Verlag, 2010.
- ²⁷Smith, A. D., G. Badlani, G. M. Preminger, and L. R. Kavoussi. Smith's Textbook of Endourology, 3rd edn. Chichester: Wiley-Blackwell, 2012.
- ²⁸Snedecker, J. G., P. Niederer, F. R. Schmidlin, M. Farshad, C. K. Demetropoulos, J. B. Lee, and K. H. Yang. Strain-rate dependent material properties of the porcine and human kidney capsule. *J. Biomech.* 38:1011–1021, 2005.
- ²⁹Spearman, C. The proof and measurement of association between two things. *Am. J. Psychol.* 15:72–101, 1904.
- ³⁰Swindle, M. M., and A. C. Smith. Comparative anatomy and physiology of the pig. *Scand J. Lab. Anim. Sci.* 25(Suppl):11–21, 1998.
- ³¹van Gerwen, D. J., J. Dankelman, and J. J. van den Dobbelaar. Needle–tissue interaction forces—a survey of experimental data. *Med. Eng. Phys.* 34:665–680, 2012.
- ³²Vidal, F. P., N. W. John, A. E. Healey, and D. A. Gould. Simulation of ultrasound guided needle puncture using patient specific data with 3d textures and volume haptics. *Comput. Anim. Virtual Worlds.* 19:111–127, 2008.
- ³³Villard, P. F., F. P. Vidal, L. ap Cenydd, R. Holbrey, S. Pisharody, S. Johnson, A. Bulpitt, N. W. John, F. Bello, and D. Gould. Interventional radiology virtual simulator for liver biopsy. *Int. J. CARS,* 2013, pp. 1–13.
- ³⁴Weibull, W. A statistical distribution function of wide applicability. *J. Appl. Mech.* 18(3):293–297, 1951.
- ³⁵Zhai, J., K. Karuppasamy, R. Zvavanjanja, M. Fisher, D. Gould, and T. How. A sensor for needle puncture force measurement during interventional radiological procedures. *Med. Eng. Phys.* 35:350–356, 2013.