



Published in final edited form as:

*Urology*. 2018 April ; 114: 114–120. doi:10.1016/j.urology.2018.01.030.

## Development and Validity of a Silicone Renal Tumor Model for Robotic Partial Nephrectomy Training

Steven M. Monda, Jonathan R. Weese, Barrett G. Anderson, Joel M. Vetter, Ramakrishna Venkatesh, Kefu Du, Gerald L. Andriole, and Robert S. Figenshau

Division of Urological Surgery, Washington University School of Medicine

### Abstract

**Objectives**—To provide a training tool to address the technical challenges of robot-assisted laparoscopic partial nephrectomy, we created silicone renal tumor models using 3D printed molds of a patient’s kidney with a mass. In this study, we assessed the face, content, and construct validity of these models.

**Materials and Methods**—Surgeons of different training levels completed four simulations on silicone renal tumor models. Participants were surveyed on the usefulness and realism of the model as a training tool. Performance was measured using operation specific metrics, self-reported operative demands (NASA TLX), and blinded expert assessment (GEARS).

**Results**—24 participants included attending urologists, endourology fellows, urology residents, and medical students. Post-training surveys of expert participants yielded mean results of 79.2 on the realism of the model’s overall feel and 90.2 on the model’s overall usefulness for training. Renal artery clamp times and GEARS scores were significantly better in surgeons further in training ( $p = 0.005$ ,  $p = 0.025$ ). Renal artery clamp times, preserved renal parenchyma, positive margins, NASA TLX, and GEARS scores were all found to improve across trials ( $p < 0.001$ ,  $p = 0.025$ ,  $p = 0.024$ ,  $p = 0.020$ ,  $p = 0.006$  respectively).

**Conclusions**—Face, content, and construct validity were demonstrated in the use of a silicone renal tumor model in a cohort of surgeons of different training levels. Expert participants deemed the model useful and realistic. Surgeons of higher training levels performed better than less experienced surgeons in various study metrics, and improvements within individuals were observed over sequential trials. Future studies should aim to assess model predictive validity, namely the association between model performance improvements and improvements in live surgery.

### Keywords

partial nephrectomy; robotic surgery; simulation training; urological surgical procedure; kidney neoplasms; laparoscopy

---

**Corresponding Author:** Steven Monda, Division of Urologic Surgery, Washington University School of Medicine, Address: 4167 Washington Boulevard, Unit A, St. Louis, MO 63108, Telephone: (509) 750-1766, Fax: (314) 367-5016, m.monda@wustl.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Objective

Every year more patients are diagnosed with localized kidney cancer, and a greater percentage of these patients are treated with robot-assisted laparoscopic partial nephrectomy (RALPN).<sup>1-3</sup> The techniques involved in RALPN continue to advance with improvements in suturing methods, robotic technology, and ischemia strategies.<sup>4-6</sup> Despite these technical advances and its increasing prominence, RALPN remains challenging with a significant learning curve. Surgeons may require as many as 20 cases before their robotic console times begin to optimize, and over 30 cases before their warm ischemia times—the duration the renal artery is clamped—begin to nadir.<sup>7</sup> This is about more than just saving minutes, renal function, and patient prognosis are both related to warm ischemia time.<sup>8</sup>

While RALPN is becoming a common procedure and a clinically meaningful learning curve has been demonstrated, there is currently no widely used training model for RALPN. Urology residents must train in the operating room on live cases with varying degrees of involvement in the actual procedure. The need to establish some of this experience outside of live surgery without compromising patient care is evident. Agarose gel injections and Styrofoam ball attachments have been described for simulating training tumors in porcine kidneys.<sup>9,10</sup> Agarose tends to extravasate while Styrofoam offers a predictable and brittle tumor margin that has a tendency to peel from the underlying kidney. Both offer limited ability to represent true tumor morphology and require porcine tissues with a limited shelf life.

Renal tumor models created by 3D printing have been described for preoperative planning in partial nephrectomy, where patient anatomy can be accurately represented and tactilely assessed.<sup>11</sup> Similarly, 3D printed silicone renal tumor models have been used for preoperative rehearsal before complex RALPN where surgeons can cut and suture on the model.<sup>12</sup> In view of this, we created silicone renal tumor models for training purposes. We implemented methods described by Cheung et al. in a pediatric pyeloplasty model to inject silicone blends into 3D printed molds of a patient's kidney with a tumor.<sup>13</sup>

In this study, we assess the validity of this silicone model as a training tool for the inexperienced surgeon. Three established types of validity for surgical models were assessed.<sup>14,15</sup> Face validity and content validity, defined as model realism and usefulness respectively, were evaluated from survey results. Construct validity, or a model's ability to distinguish between trainees of different experience levels, was assessed using various metrics of simulation performance which were compared across training levels and trials. Metrics were gathered from simulation specific measurements (e.g. renal artery clamp time), self-assessment, and blinded expert review.

## Materials and Methods

### 3D Model Generation

A representative kidney with a single tumor was selected from MRI imaging of a patient who underwent RALPN at our institution. The tumor was deemed of medium complexity

and good for training purposes. It was four centimeters in diameter, moderately exophytic, within four millimeters of the hilar vessels, and midline. The assigned R.E.N.A.L nephrometry score was eight.<sup>18</sup>

The kidney and tumor were rendered into a 3D format from MRI imaging using the open source segmentation software, Invesalius (Centro da Tecnologia da informação, Amarais, Brazil). Mold files (STL file format) were then created using the open source 3D modeling software Blender (Blender Foundation, Amsterdam, Netherlands). Negative volume molds were created of both the kidney with the tumor removed and the kidney with the tumor in place to allow for a two-step molding process. 3D printing of molds was then performed on an Object Eden260VS printer (Stratasys, Eden Prairie, MN). Molds could then be used for repeated castings.

Models were cast with a blend of Dragonskin® 20 silicone and Slacker® silicone deadener (Smooth-On, Inc., Macungie, PA) using methods described for the development of a pediatric pyeloplasty model.<sup>13</sup> A mass ratio of 9:4, silicone to deadener, was mixed and then injected using a syringe into the kidney *without tumor* mold. After curing, the resulting solid kidney was placed in the kidney *with tumor* mold. Silicone was then injected into the tumor cavity. Silicone pigment was used to color tumor and kidney appropriately during each step. After the tumor silicone cured, the model was removed, and a section of dyed surgical tubing was stitched at the renal hilum to simulate the renal artery for bulldog clamp application. Each completed silicone kidney with tumor used 125 grams of silicone material at a cost of \$3.90 USD and required around 15 minutes of active labor.

## Participants

Medical students, urology residents, endourology fellows, and attending urologists were recruited from our institution and consented to participate in this study. All medical students had completed a urology rotation and shadowed a partial nephrectomy prior to recruitment. Less focus was placed on medical student recruitment as they were deemed less representative trainees. Participants were categorized based on their training levels which were expected to best reflect levels of robotic surgery experience at our institution (Table 1). This study was approved by our Institutional Review Board.

## Surgical Simulation

Each participant performed two simulated surgeries on two different days for four total trials. A minimum of one week was interposed between trials two and three. The same written script was used for instruction for all participants. These trial surgeries were performed on a da Vinci Si surgical robot (Intuitive Surgical, Sunnyvale, CA) with the model placed in a laparoscopic trainer box. Steps incorporated into each simulation were instrument choice, port placement on trainer box, intraoperative ultrasound evaluation, clamping the renal artery, tumor excision with retraction, capsular renorrhaphy, and specimen entrapment. These steps were selected based on their clinical relevance, associated difficulty, and feasibility of incorporation. Tumor excision with retraction and capsular suturing were considered critical to surgical outcomes and difficult skills to acquire.

All participants received the same instructions prior to surgery and were encouraged to treat this as a live case. The same bedside assistant, who through a laparoscopic port handed off sutures, placed clips, removed needles, changed robotic instruments, etc., was used for all trials. Subjects were notified of their warm ischemia time and could appreciate their resection specimen after each trial. All trials were recorded from the da Vinci video output.

### Metrics Assessed

The three operation specific metrics (renal artery clamp time, preserved renal parenchyma, and surgical margins) were all assessed by the same researcher who was blinded to participant and trial number. Renal artery clamp time was measured from video recording as the time from initial bulldog clamp placement on the renal artery to clamp removal. Subjects were allowed to unclamp when they felt they had achieved an adequately hemostatic closure. Preserved renal parenchyma was taken as the ratio of the final kidney mass *after specimen excision* to the initial kidney mass *without tumor*, as measured during the molding process. Surgical margins were assessed by sectioning the resection specimen. Both clamp time and preserved renal parenchyma are related to long-term kidney functions, and negative surgical margins are necessary for complete oncologic resection.<sup>8,19</sup>

Face and content validity were assessed after participants completed their last trials using survey questions with a 0–100 slider bar scale anchored at unrealistic-realistic and useless-useful respectively. Trainee self-assessed operative demand was surveyed immediately after the completion of each trial using NASA Task Load Index (NASA TLX).<sup>16</sup> Each subscale of NASA TLX was evaluated on a 0–100 point slider bar scale using established anchors. NASA TLX is regularly used for assessing task demand in surgery.<sup>20</sup>

Blinded expert-assessed surgical performance was measured using the standardized and validated Global Evaluative Assessment of Robotic Surgeons (GEARS).<sup>21</sup> Two fellowship trained attending surgeons, who have performed over 200 RALPN each, evaluated the first and fourth trial of all participants. Evaluations were based on video recordings of the robotic display without audio. Videos were deidentified and distributed in a random order. Evaluators were thus blinded to both trainee and trial number. Each evaluator unknowingly reviewed the first and fourth trials from the same trainees to ensure that differences within participants could be properly assessed.

The Kruskal-Wallis test was used for quantitative demographics comparisons between experience levels. Multilevel multivariable linear and logistic regression models were used to test the association of performance metrics against trial number and experience level. Performance metrics were used as the dependent variables. Trial number and experience level were used as independent variable. Each individual was treated as a random effect. Significance levels were set at 0.05.

## Results

### Participants

A total of 24 participants were recruited for this study including four medical students, 14 residents, three endourology fellows, and three attending surgeons. Participant demographics

and experience levels are listed in Table 1. Participants were grouped as follows, medical students and interns, second and third year residents, fourth and fifth year residents, and fellows and attendings.

### Face and content validity

Face and content validity results from our study's expert participants—endourology fellows and attending urologists—are displayed in Figure 1. Mean responses of 79.2 were achieved in the realism of the model's overall feel and 90.7 in the model's overall usefulness for training. The model was deemed most useful, “for trainees to obtain new technical skills” with a mean score of 93.8. The lowest mean usefulness was recorded “for trainees to improve existing technical skills” at 85.7. The highest metrics for model realism were achieved for needle driving, cutting, and visual representation at 78.3, 78.0, and 78.0. The lowest metrics of model realism, kidney malleability and tumor-kidney differentiation, achieved mean scores of 67.0 and 74.7

### Construct Validity

Construct validity was assessed with *within individuals* and *between groups* comparing various performance metrics across trials and between experience levels. Operation specific metrics, certain NASA TLX scores, and GEARS scores were all found to improve across trials one through four ( $p = 0.025$ ,  $p = 0.020$ ,  $p = 0.006$ ) (Table 2). Only the NASA Task Load Index item, Physical Demand was found to not improve across trials ( $p=0.130$ ). Renal artery clamp times were found to be higher in advanced and expert surgeons compared to novice surgeons ( $p=0.005$ ,  $p<0.001$ ) (Table 3). GEARS scores were also found to be improved in advanced and expert surgeons in all GEARS metrics ( $p = 0.025$ ,  $p = 0.004$ ). No differences were observed between intermediate and novice participants across all metrics ( $p = 0.20$ ).

### Discussion

RALPN has seen increased adoption with little corresponding progress in how the procedure is taught. As simulation becomes a mainstay of urology resident education, it is expected to fill a gap for this complex procedure. The ability to gain surgical skills in a stress-free environment without compromising live patients has appeal in RALPN where tumor margins and warm ischemia times add a level of urgency not seen in other urological procedures. This study aims to provide a preliminary assessment of the internal validity of a silicone renal tumor model for RALPN training

The perceived advantages of our approach are many. 3D printing allows for accurate representation of patient anatomy from MRI or CT images. Molding provides a unique opportunity to quickly produce many durable kidneys reliably.<sup>22,23</sup> Silicone is suggested to represent kidney tissue in terms of tear strength, and 3D printed silicone kidneys are demonstrated to be useful for preoperative rehearsal.<sup>12,22,23</sup> Finally silicone has been used elsewhere in surgical simulation for representing ureters, paranasal sinuses, cerebral aneurysms, and most commonly, skin in suture pads.<sup>13,24,25</sup>

Our surveys of model realism established preliminary face validity. Notably, high measures of realism were achieved for both needle driving and cutting (means 78.3 and 78.0) as well

as overall feel (mean 79.2). Our model also achieved content validity with expert participants deeming the model useful (Fig. 1). The highest survey measures of usefulness were found for trainees to obtain new technical skills, followed by trainee evaluation, and then existing skill improvement (means 93.8, 87.7, and 85.7 respectively). These results suggest that this model's most meaningful role may be to introduce naive trainees to RALPN while still providing a means for more advanced trainees to refine existing skills. As model fidelity improves, we expect that the model's usefulness for more experienced trainees will also improve.

Further research should assess and aim to improve our current model realism, specifically kidney malleability which had the lowest face validity (mean 67.0). Objective measures of model tear strength, hardness, and resistance to needle driving and cutting would supplement both our current subjective face validity surveys and past research to more conclusively assess model realism.<sup>22,23</sup>

Other partial nephrectomy simulations have incorporated porcine kidneys and simulated perfusion, neither of which were used in our model. Porcine kidneys may represent human tissue more closely, and managing bleeding is a critical aspect of surgery. This must be weighed against the advantages of a simple, synthetic, economic, easy-to-setup model that has practically indefinite shelf life and can accurately represent diverse tumor geometries. Given the sporadic and busy nature of residency, having models readily available has evident advantages. Furthermore, synthetic models have provided more cost-effective means of training novice surgeons compared to animal tissue models.<sup>26,27</sup>

Intraoperative ultrasound was used in all simulations. However, there was some limitation in the ability to differentiate kidney and tumor sonographically as both tumor and kidney had limited echogenicity. Recently methylcellulose is suggested to enhance ultrasound silicone echogenicity, and future work will establish silicone additives to allow proper ultrasound evaluation of tumor margins.

Crucially, our model demonstrated construct validity, or the ability to distinguish between different levels of experience both *between groups* and *within individuals*.<sup>14</sup> More experienced surgeons performed better than less experienced surgeons. Significant differences were found between more and less experienced surgeons in renal artery clamp times and GEARS scores (Table 3). This suggests that the model could distinguish more proficient surgeons based on their simulation performance. Our model also demonstrated construct validity *within individuals* where trainee performance was found to improve across trials. Except physical demand, all study metrics were found to improve across trials ( $p < 0.05$ ), see Table 2. Given the already low physical demand of robotic surgery in short durations, it is understandable that reductions were not observed.

There were other limitations in assessing differences between training levels. The significant variability in how individuals report operative demand precluded significant differences between groups in NASA TLX. The consistent results in preserved renal parenchyma across subjects likely explains the lack of significant differences between groups in this metric. It is unclear why positive margins were not found to differ between groups. Significant

differences were also not found between our novice and intermediate participants which likely reflects their overall inexperience (Table 1). Although the intermediate participants had assisted more surgeries than the novice group, no participants from either group had completed a RALPN independently.

The economic nature of this model makes it feasible to implement in residency training programs. Silicone injection provides a means for producing reliable surgical models quickly and can be performed with readily available research equipment - syringes, a scale, containers, and 3D printers for initially making molds as has been described throughout the surgical literature.<sup>13,24,25</sup> The 125 grams of nontoxic silicone used in each of our models had a material cost of \$3.90 USD online. The initial 3D printing of two molds for silicone injection cost \$260 USD at our institution. Most hospitals and research institutions now have access to 3D printers, and most 3D printers work for printing molds with higher resolution printers resulting in more accurate molds. However, access to 3D printing does potentially represent a financially significant requirement for the development of these models. The injection and curing of silicone models required 20 minutes of active injection and mixing followed by 1.5 hours of silicone curing. Faster cure silicones and using multiple molds in parallel allows for the batch production of many models with limited labor.

Our study used one medium-difficulty midline anterior tumor across trials. This same tumor excision was repeated four times per trainee. The decision to use the same renal tumor was made to ensure that construct validity *within individuals* could be established through improvements across trials. This would be difficult to assess if tumor difficulty was variable. Future work will certainly establish different kidney tumors and thus provide advancing training difficulty. Large endophytic tumors and tumors with complex borders as well as easier exophytic tumors are all possible using 3D printed molds.

Surgery is an intricate and nuanced skill with complexities that no simulation can expect to fully encompass. We have explored a silicone renal tumor model with the primary goal of familiarizing urological trainees with key aspects of RALPN. Nevertheless, many nuances of RALPN were not addressed. Bowl mobilization, landmark identification, kidney exposure, hilar dissection, management of urine leakage and bleeding, and pelviciceal reconstruction are all key aspects of RALPN that were not incorporated into our simulation. The aspects of RALPN that were incorporated were based on feasibility and perceived importance.

The most meaningful measure of any surgical training model is the extent to which model performance is associated with operating room performance improvements. Such predictive validity is difficult to assess and is not often established,<sup>14</sup> yet it is essential for assessing whether a training model is achieving its desired effect, namely better operating room performance by the trainee and better outcomes for the patient. Future research should assess the association between warm ischemia times and GEARS scores in live surgery with model use and performance. This predictive validity would complement the face, content, and construct validities that have begun to be established in this study.

## Conclusion

This study provides an initial assessment of a silicone renal tumor model for RALPN training. It complements current literature on the use of silicone renal tumor models for preoperative rehearsal. In describing our novel methods of manufacturing silicone models, we suggest an economic means for trainees to gain fundamental surgical skills in RALPN.

Our model achieved three important levels of validation. Face and content validities were established using survey assessments of model realism and usefulness. Construct validity *between groups* was demonstrated with more experienced surgeons achieving better metrics of performance. Construct validity *within individuals* was also established demonstrating improvements over the course of multiple trials. Further research will establish tumor models of different difficulties. Additionally, future work will assess predictive validity looking at associations between trainee performance on silicone models and improvements in trainee performance in live surgery.

## Supplementary Material

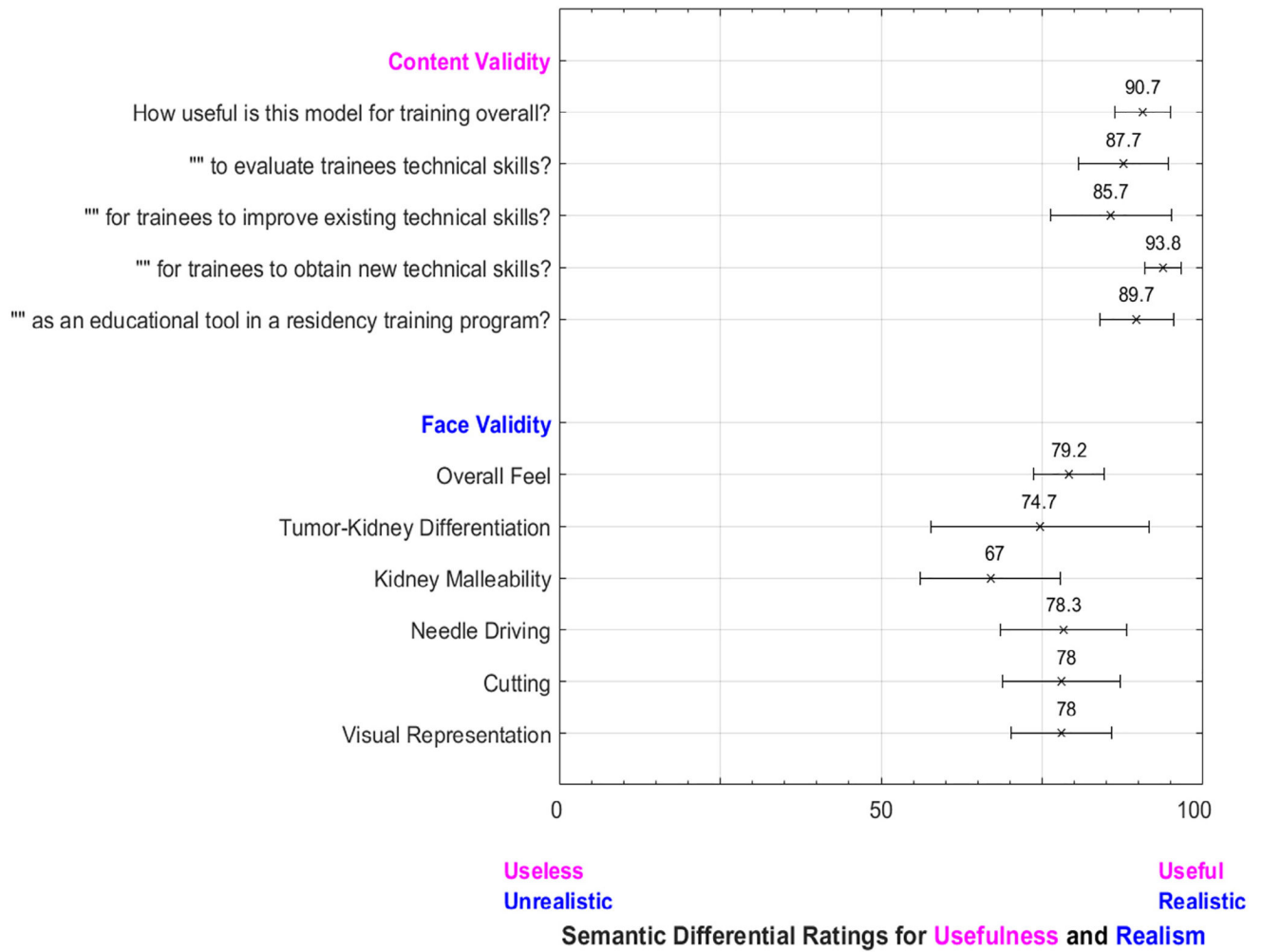
Refer to Web version on PubMed Central for supplementary material.

## References

1. SEER Cancer Stat Facts: Kidney and Renal Pelvis Cancer. National Cancer Institute; Bethesda, MD: <https://seer-cancer.gov.beckerproxy.wustl.edu/statfacts/html/kidrp.html> [Accessed July 10, 2017]
2. Gandaglia G, Ravi P, Abdollah F, et al. Contemporary incidence and mortality rates of kidney cancer in the United States. *Can Urol Assoc J.* 2014; 8(7–8):247–252. DOI: 10.5489/auj.1760 [PubMed: 25210548]
3. Liu J-J, Leppert JT, Maxwell BG, Panousis P, Chung BI. Trends and perioperative outcomes for laparoscopic and robotic nephrectomy using the National Surgical Quality Improvement Program (NSQIP). *Urol Oncol Semin Orig Investig.* 2014; 32(4):473–479. DOI: 10.1016/j.urolonc.2013.09.012
4. Sammon J, Petros F, Sukumar S, et al. Barbed suture for renorrhaphy during robot-assisted partial nephrectomy. *J Endourol.* 2011; 25(3):529–533. DOI: 10.1089/end.2010.0455 [PubMed: 21351886]
5. Aboumarzouk OM, Stein RJ, Eyraud R, et al. Robotic Versus Laparoscopic Partial Nephrectomy: A Systematic Review and Meta-Analysis. *Eur Urol.* 2012; 62(6):1023–1033. DOI: 10.1016/j.eururo.2012.06.038 [PubMed: 22771266]
6. Funahashi Y, Yoshino Y, Sassa N, Matsukawa Y, Takai S, Gotoh M. Comparison of Warm and Cold Ischemia on Renal Function After Partial Nephrectomy. *Urology.* 2014; 84(6):1408–1413. DOI: 10.1016/j.urology.2014.08.040 [PubMed: 25432829]
7. Mottrie A, De Naeyer G, Schatteman P, Carpentier P, Sangalli M, Ficarra V. Impact of the Learning Curve on Perioperative Outcomes in Patients Who Underwent Robotic Partial Nephrectomy for Parenchymal Renal Tumours. *Eur Urol.* 2010; 58(1):127–133. DOI: 10.1016/j.eururo.2010.03.045 [PubMed: 20399002]
8. Volpe A, Blute ML, Ficarra V, et al. Renal Ischemia and Function After Partial Nephrectomy: A Collaborative Review of the Literature. *Eur Urol.* 2015; 68(1):61–74. DOI: 10.1016/j.eururo.2015.01.025 [PubMed: 25703575]
9. Hung AJ, Ng CK, Patil MB, et al. Validation of a novel robotic-assisted partial nephrectomy surgical training model. *BJU Int.* 2012; 110(6):870–874. DOI: 10.1111/j.1464-410X.2012.10953.x [PubMed: 22313582]



10. Taylor GD, Johnson DB, Hogg DC, Cadeddu JA. Development of a renal tumor mimic model for learning minimally invasive nephron sparing surgical techniques. *J Urol.* 2004; 172(1):382–385. DOI: 10.1097/01.ju.0000132358.82641.10 [PubMed: 15201816]
11. Wake N, Rude T, Kang SK, et al. 3D printed renal cancer models derived from MRI data: application in pre-surgical planning. *Abdom Radiol.* 2017; 42(5):1501–1509. DOI: 10.1007/s00261-016-1022-2
12. von Rundstedt F-C, Scovell JM, Agrawal S, Zaneveld J, Link RE. Utility of patient-specific silicone renal models for planning and rehearsal of complex tumour resections prior to robot-assisted laparoscopic partial nephrectomy. *BJU Int.* 2017; 119(4):598–604. DOI: 10.1111/bju.13712 [PubMed: 27862866]
13. Cheung CL, Looi T, Lendvay TS, Drake JM, Farhat WA. Use of 3-Dimensional Printing Technology and Silicone Modeling in Surgical Simulation: Development and Face Validation in Pediatric Laparoscopic Pyeloplasty. *J Surg Educ.* 2014; 71:762–767. DOI: 10.1016/j.jsurg.2014.03.001 [PubMed: 24776857]
14. Schout BMA, Hendriks AJM, Scherpier AJJA, Bemelmans BLH. Update on Training Models in Endourology: A Qualitative Systematic Review of the Literature between January 1980 and April 2008. *Eur Urol.* 2008; 54(6):1247–1261. DOI: 10.1016/j.eururo.2008.06.036 [PubMed: 18597924]
15. McDougall EM. Validation of Surgical Simulators. *J Endourol.* 2007; 21(3)doi: 10.1089/end.2007.9985
16. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. 1988; :139–183. DOI: 10.1016/S0166-4115(08)62386-9
17. Goh AC, Goldfarb DW, Sander JC, Miles BJ, Dunkin BJ. Global Evaluative Assessment of Robotic Skills: Validation of a Clinical Assessment Tool to Measure Robotic Surgical Skills. *J Urol.* 2012; 187(1):247–252. DOI: 10.1016/j.juro.2011.09.032 [PubMed: 22099993]
18. Kutikov A, Uzzo RG. The R.E.N.A.L. Nephrometry Score: A Comprehensive Standardized System for Quantitating Renal Tumor Size, Location and Depth. *J Urol.* 2009; 182(3):844–853. DOI: 10.1016/j.juro.2009.05.035 [PubMed: 19616235]
19. Mir MC, Ercole C, Takagi T, et al. Decline in Renal Function after Partial Nephrectomy: Etiology and Prevention. *J Urol.* 2015; 193(6):1889–1898. DOI: 10.1016/j.juro.2015.01.093 [PubMed: 25637858]
20. Zheng B, Jiang X, Tien G, Meneghetti A, Panton ONM, Atkins MS. Workload assessment of surgeons: correlation between NASA TLX and blinks. *Surg Endosc.* 2012; 26(10):2746–2750. DOI: 10.1007/s00464-012-2268-6 [PubMed: 22527300]
21. Aghazadeh MA, Jayaratna IS, Hung AJ, et al. External validation of Global Evaluative Assessment of Robotic Skills (GEARS). *Surg Endosc.* 2015; 29(11):3261–3266. DOI: 10.1007/s00464-015-4070-8 [PubMed: 25609318]
22. Alamri AA, Abdulla A, Madjeruh J, et al. Designing A High-Fidelity Laparoscopic Partial Nephrectomy Bench Model: Determining the Tear Strength and Resistance of a Synthetic Silicone Composition. *J Urol.* 2011; 185(4):e597–e598. DOI: 10.1016/j.juro.2011.02.1449
23. Alamri A, Abdulla A, Madjeruh J, Matsumoto ED. Validation of a Partial Nephrectomy Bench Model Developed Via a Novel Material Engineering Process. *J Urol.* 2012; 187(4):e608.doi: 10.1016/j.juro.2012.02.1270
24. Chang DR, Lin RP, Bowe S, et al. Fabrication and validation of a low-cost, medium-fidelity silicone injection molded endoscopic sinus surgery simulation model. *Laryngoscope.* 2017; 127(4):781–786. DOI: 10.1002/lary.26370 [PubMed: 28000224]
25. Ryan JR, Almefty KK, Nakaji P, Frakes DH. Cerebral Aneurysm Clipping Surgery Simulation Using Patient-Specific 3D Printing and Silicone Casting. *World Neurosurg.* 2016; 88:175–181. DOI: 10.1016/j.wneu.2015.12.102 [PubMed: 26805698]
26. Grober ED, Hamstra SJ, Wanzel KR, et al. The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures. *Ann Surg.* 2004; 240(2):374–381. DOI: 10.1097/01.sla.0000133346.07434.30 [PubMed: 15273564]
27. Matsumoto ED, Hamstra SJ, Radomski SB, Cusimano MD. The Effect of Bench Model Fidelity On Endourological Skills: A Randomized Controlled Study. *J Urol.* 2002; 167(3):1243–1247. DOI: 10.1016/S0022-5347(05)65274-3 [PubMed: 11832706]



**Figure 1.** Face and content validity survey items and results. Mean results from expert participants (n=6) are displayed with 95% confidence intervals.

**Table 1**

Participant demographics and experience (means).

Experience Level*	Beginner				Intermediate				Advanced				Expert	
	Medical Student	Intern	URO-1	URO-2	URO-3	URO-4	URO-1	URO-2	URO-3	URO-4	Fellow	Attending		
n	4	2	3	3	2	4	3	3	2	4	3	3		
Age, years	25.8	26.0	28.0	29.0	31.0	31.3	31.3	32.3	32.3	32.3	32.3	47.7		
Experience performing partial nephrectomy, years	0	0	0	0.5	1.5	2.5	2.5	2.7	2.7	2.7	2.7	16.3		
RALPN performed (completed >50% of the case), cases	0	0	0	0	0	4.8	16.3	168.3	168.3	168.3	168.3	168.3		
RALPN assisted (completed <50% of the case), cases	0	0	1.0	7.0	27.5	32.5	40.0	33.3	33.3	40.0	40.0	33.3		
All robotic surgeries performed or assisted, cases	0.3	5.0	3.3	60.0	82.5	110.0	183.3	561.7	561.7	183.3	183.3	561.7		

\* Comparisons made across experience levels using Kruskal-Wallis Test yielded p-values <0.001 for all listed variables.

**Table 2**

Mean trainee performance across trials and results of linear regression analysis.

<u>Performance Metrics*</u>	Trial 1	Trial 2	Trial 3	Trial 4	Estimate (95% CI)	P-value*
<b>Operation Specific Metrics</b>						
Clamp Times (minutes)	20.5	16.2	14.6	13.8	-2.2 (-2.9, -1.4)	<0.001
Positive Margins (%)**	29%	29%	21%	4%	0.53 (0.29,0.89)	0.024
Proportion Renal Parenchyma Preserved (grams/grams)	0.98	0.97	0.97	0.97	-0.002 (-0.0038, -0.0003)	0.025
<b>NASA Task Load Index (0-100)</b>						
Mental Demand	65.5	54.3	54.5	43.7	-6.7 (-8.9, -4.5)	<0.001
Physical Demand	42.4	42.0	39.1	37.4	-1.8 (-4.0, 0.5)	0.130
Temporal Demand	57.5	51.7	47.5	41.4	-5.2 (-7.5, -3.0)	<0.001
Performance (Failure = 100)	45.5	39.2	35.8	31.2	-4.6 (-7.4, -1.9)	<0.001
Hard Work	62.5	55.2	55.0	49.1	-4.0 (-7.2, -0.9)	0.014
Frustration	43.3	42.3	38.3	33.0	-3.5 (-6.4, -0.6)	0.020
<b>GEARS Score (0-5)</b>						
Depth Perception	3.33	-	-	4.08	0.8 (0.4, 1.1)	0.001
Bimanual Dexterity	2.96	-	-	3.79	0.8 (0.5, 1.1)	<0.001
Efficiency	3.04	-	-	3.79	0.8 (0.4, 1.1)	<0.001
Force Sensitivity	2.88	-	-	3.50	0.6 (0.2, 1.0)	0.006
Autonomy	3.50	-	-	4.04	0.5 (0.5, 0.8)	<0.001
Robotic Control	2.92	-	-	3.83	0.9 (0.9, 1.2)	<0.001

\* Multilevel multivariable linear and logistic regression models were used to test the association of performance metrics against the independent variables, trial number and experience level.

\*\* Positive margins represent the only binary variable, and the "Estimate" column is the odds ratio from logistic regression in this case.

**Table 3**

Trainee performance assessed by multilevel multivariable regression. NASA-TLX comparisons are omitted from this table as there were no significant differences between experience levels – see Discussion.

	Intermediate vs. Novice	Advanced vs. Novice	Expert vs. Novice
	Estimated difference from regression analysis (95% CI) p-value		
<b>Operation Specific Metrics</b>			
Clamp Times (minutes)	-1.9 (-7.0, 3.2) 0.50	-8.8 (-13.9, -3.7) 0.005	-12.4 (17.6, -7.3) <0.001
Positive Margins **	1.02 (0.12, 8.5) 0.99	0.54 (0.05, 4.5) 0.54	0.50 (0.04, 4.0) 0.50
Proportion Renal Parenchyma Preserved	0.001 (-0.012, 0.014) 0.91	-0.010 (-0.023, 0.004) 0.19	-0.010 (-0.024, 0.003) 0.16
<b>GEARS Score</b>			
Depth Perception	-0.17 (-0.94, 0.61) 0.69	1.3 (0.56, 2.1) 0.004	1.3 (0.56, 2.1) 0.004
Bimanual Dexterity	-0.083 (-0.74, 0.57) 0.82	1.1 (0.43, 1.7) 0.006	1.5 (0.84, 2.2) <0.001
Efficiency	-0.50 (-1.20, 0.20) 0.20	0.92 (0.21, 1.6) 0.025	1.3 (0.55, 2.0) 0.004
Force Sensitivity	-0.083 (-0.66, 0.49) 0.82	1.6 (1.0, 2.2) <0.001	1.9 (1.3, 2.5) <0.001
Autonomy	-0.083 (-0.28, 0.12) 0.75	1.2 (0.97, 1.4) <0.001	1.7 (1.5, 1.9) <0.001
Robotic Control	-0.25 (-0.54, 0.056) 0.46	0.92 (0.62, 1.2) 0.012	1.2 (0.87, 1.5) 0.002

\* Multilevel multivariable linear and logistic regression models were used to test the association of performance metrics against the independent variables, trial number and experience level.

\*\* Positive margins represent the only binary variable, and the “Estimate” column is the odds ratio from logistic regression in this case.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript