Check for updates

ORIGINAL ARTICLE

The International Journal of **Medical Robotics** and **Computer Assisted Surgery**

WILEY

Analysis of the video motion tracking system "Kinovea" to assess surgical movements during robot-assisted radical prostatectomy

Alexander J. W. Beulens[1,2] 🔘 | Hanae F. Namba[2,3] | Willem M. Brinkman[4] |
Richard P. Meijer[4] | Evert L. Koldewijn[2] | Ad J. M. Hendrikx[5] |
Jean-Paul van Basten[6] | Jeroen J. G. van Merriënboer[7] | Henk G. Van der Poel[8] |
Chris Bangma[9] | Cordula Wagner[1,10]

[1]Netherlands Institute for Health Services Research (NIVEL), Utrecht, The Netherlands

[2]Department of Urology, Catharina Hospital, Eindhoven, The Netherlands

[3]Faculty of Medicine, Utrecht University, Utrecht, The Netherlands

[4]Department of Oncological Urology, University Medical Centre Utrecht, Utrecht, The Netherlands

[5]Urologist N.P., The Netherlands

[6]Department of Urology, Canisius Wilhelmina Hospital, Nijmegen, The Netherlands

[7]School of Health Professions Education, Maastricht University, Maastricht, The Netherlands

[8]Department of Urology, Dutch Cancer Institute-Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands

[9]Department of Urology, Erasmus Medical Centre, Rotterdam, The Netherlands

[10]Amsterdam Public Health Research Institute, Amsterdam UMC, Amsterdam, The Netherlands

Correspondence
Alexander J. W. Beulens, Department of Urology, Catharina Hospital, P.O. Box 1350, 5602 ZA, Eindhoven, The Netherlands.
Email: alexander.beulens@ catharinaziekenhuis.nl

Abstract

Backgrounds: Robot-assisted surgery facilitated the possibility to evaluate the surgeon's skills by recording and evaluating the robot surgical images. The aim of this study was to investigate the possibility of using a computer programme (Kinovea) for objective assessment of surgical movements in previously recorded in existing robot-assisted radical prostatectomy (RARP) videos.

Methods: Twelve entire RARP procedures were analysed by a trained researcher using the computer programme "Kinovea" to perform semi-automated assessment of surgical movements.

Results: Data analysis showed Kinovea was on average able to automatically assess only 22% of the total surgical duration per video of the robot-assisted surgery. On average, it lasted 4 hours of continued monitoring by the researcher to assess one RARP using Kinovea.

Conclusion: Although we proved it is technically possible to use the Kinovea system in retrospective analysis of surgical movement in robot-assisted surgery, the acquired data do not give a comprehensive enough analysis of the video to be used in skills assessment.

KEYWORDS

video analysis, surgical skills, surgical movements, robot assisted surgery, robot assisted radical prostatectomy

1 | INTRODUCTION

The introduction of endoscopic surgery opened the possibility to evaluate the surgeon's skills based on the intra-corporal surgical videos. A standardized assessment method in order to assess the surgeon's skills real-time in surgery was developed by Martin et al using the objective structured assessment of technical skills (OSATS).[1] This type of skill assessment sparked the investigation into the relation between surgical skills and post-operative complications. Birkmeyer et al were among the first who were able to prove a causal relation between the level of surgical skill[2] and post-operative complications. They demonstrated that the skills of

the surgeons in laparoscopic bariatric surgery were associated with lower complication and mortality rates. Moreover, lower ranked surgical skills were associated with prolonged surgical times, higher reoperation rates and higher readmission rates.[2]

After the initial development of OSATS, various other assessment tools have been developed for the evaluation of robotic surgical technical skills, such as the global evaluative assessment of robotic skills (GEARS),[3] the prostatectomy assessment and competency evaluation (PACE),[4] the generic error rating tool (GERT),[5] and the PROTEST assessment tools.[6] These new assessment tools can be used to assess the surgeon's skills based on the intra-corporal video of the surgical procedure.

In another study conducted by Goldenberg et al, the relationship between surgical skills defined by use of the GEARS assessment tool and patient outcomes in robot-assisted radical prostatectomies (RARPs) was investigated.[7] The focus of the GEARS assessment tool lies on general robot surgical principles, that is, depth perception, bimanual dexterity, efficiency, force sensitivity, autonomy and robotic control.[3] Goldenberg et al found that surgical skills as measured using the GEARS tool were ranked higher in the post-operative urinary continent group compared to post-operative urinary incontinent group. However, these assessment tools are a time-consuming way of assessing surgeons' skills. Since these methods are based on human review, subjective bias cannot be avoided, leading to reduced inter-rater reliability.

To objectify the assessment and to reduce the time investment of "manual assessment" by the observer automated assessment tools are being studied.[8,9] Recently Hung et al investigated the possibility of the analysis of the movements of the surgical robot with the dVLogger system which led to greater insight into the performance of the surgeons and could be used to predict post-operative outcomes.[9,10] The dVLlogger system automatically logs motion tracking and system events data without correlation with the surgery.

In a recent Delphi study by our group[6] into the link between surgical skills and post-operative outcome (resulting in the PROTEST assessment tool) a possible relation was suggested between some phases of the RARP surgery and post-operative urinary continence. The Delphi panel agreed based on expert opinion that events during the "pelvic floor muscle exposure/opening of the endopelvic fascia," and the "urethrovesical anastomosis" could be related to post-operative continence. This relation has yet to be proven in further research.

Ganni et al researched a computer programme, Kinovea, which can be used by researchers to perform semi-automated video motion tracking in laparoscopic cholecystectomy procedures. The system is able to track pixels selected by the researcher on the surgical instrument during the surgical procedure. It follows the position of the pixel during the procedure, which enables the software to track the position of the surgical instrument. Kinovea is a software-based system used in sports to track trajectories and speeds of items or human movements. The benefit of this system over the dVlogger is that there is no need of additional hardware

systems or sensors on the instruments to measure the movements of the item of interest. This means this system can be used in retrospective studies without any preparation during surgery. The system enables the assessment of the video material during the tracking analysis since tracking is done in the video itself. This tracking system was used to assess surgeons' skills using existing surgical videos rather than in a simulator.[8] This raises the question if Kinovea could be used for similar purposes in robot-assisted surgery in order to assess the surgical movements based on the video of the surgical procedure.

The purpose of this study is to determine whether Kinovea is a valid tool for automated surgical movements tracking in RARP and may be used to evaluate a possible relation between surgical movements and post-operative urinary continence in RARP.

The present research questions are: (a) is Kinovea a valid tool for automated assessment of surgical movements in RARP surgical videos? (b) Can the results found through automated surgical movements tracking using Kinovea be used to predict post-operative continence in RARP? And (c) can results of the Kinovea analysis obtained during the "pelvic floor muscle exposure/opening of the endopelvic fascia," and the "urethrovesical anastomosis" be used to predict post-operative continence in RARP? These questions will be answered by analysing surgical movements in RARP videos.

2 | METHODS

2.1 | Subjects

For this study, existing videos of RARP procedure were used of patients who underwent a RARP in a specialized cancer hospital in the Netherlands between June 2009 and February 2017. All of the patients were operated by the same expert robotic surgeon Henk van der Poel (HvdP), who had performed >220 RARP procedures before June 2009 and has currently performed >2100 RARP procedures using the da Vinci Si surgical robot by Intuitive. Exclusion criteria were urinary incontinence prior to surgery and surgeries where no or incomplete video material was available. All of the men included had localized prostate cancer (cT1c-cT3a, Nx-N0, Mx-M0).

2.2 | Design

The study design was a pilot study in order to investigate whether Kinovea is suitable to use in robot-assisted surgery. Patient results were obtained prospectively. The follow-up was at least 12 months.

Cases were anonymized and labelled with study codes, meaning the researcher was blinded to all patient characteristics and outcomes. This study was granted approval from the institutional medical committee.

2.3 | Automated surgical movements assessment using Kinovea

The surgical movements of the instrument were tracked using the Kinovea software. The primary outcome measurements are the total time analysed (min) and percentage of surgery analysed (%). The secondary outcome measurements are total path length (cm), number of sudden movements (defined as more than 1 cm movement of the instrument per frame of the surgical video), and average speed (cm/s). The Kinovea software is deemed valid if it is able to track >80% of the duration of the surgery.

The instrument controlled by the right robotic arm (controlled by the dominant hand of the surgeon) was used for the analysis using Kinovea. For every video, excel sheets containing automatic calculations of the total distances and velocities per trajectory were downloaded from Kinovea. These results were compared to the total distances, average velocities and numbers of sudden movements calculated using manual formulas based on Ganni et al's article (Table 1).[8]

2.4 | Relation between Kinovea results and post-operative continence status

Although patient selection was based on continence, a number of additional outcomes were compared to the video motion tracking data in order to assess whether there was a relation. These post-operative outcomes included patient reported outcome measures (PROMs), lower urinary tract symptoms, measured using the International Prostate Symptom Score (IPSS),[11,12] the lower urinary tract symptoms domain of the EORTC QLQ-PR25 score (EORTC QLQ-PR25 score),[11,13,14] and post-operative complications, which were registered in the patients' medical files as they occurred.

TABLE 1 Formulas used for manual calculation[9]

Description	Symbol	Formula
Number of frames	N	
Number of frames per second	f	
Scaled x-coordinate at frame n (cm)	x'_n	
Scaled y-coordinate at frame n (cm)	y'_n	
Distance moved between consecutive frames (cm)	d_{n+1}	$\sqrt{\left(x'_{n+1}-x'_n\right)^2 + \left(y'_{n+1}-y'_n\right)^2}$
Total time taken (s)	T	$\frac{N}{f}$
Total distance travelled (cm)	D	$\sum\limits_{n=1}^{N-1} d_{n+1}$
Average speed (cm/s)	$\bar{S}$	$\frac{D}{T}$
Number of sudden movements	M^e_{n+1}	$\begin{cases} 1 & d_{n+1} \geq d_{th} \\ 0 & d_{n+1} < d_{th} \end{cases}$
Number of sudden movements	E	$\sum\limits_{n=1}^{N-1} M^e_{n+1}$

2.5 | Relation between Kinovea results of different phases of the surgery and post-operative continence status

The surgery was divided in seven surgical phases which were defined in the PROTEST assessment method[6] developed by this research group. To investigate the relation between Kinovea results and the "pelvic floor muscle exposure/opening of the endopelvic fascia," and "urethrovesical anastomosis" specific analysis for these phases were performed.

2.6 | Methods of measurement

The automated surgical movements tracking was performed by one researcher using Kinovea 0.8.15. The researcher was instructed in the use of Kinovea and the surgical procedure by a researcher experienced in the use of Kinovea and the RARP procedure.

The tracked instrument controlled by the right robotic arm in the majority of measurements was the Intuitive Surgical Hot Shears (Monopolar Curved Scissors). This instrument had a jaw length of 1.3 cm, which was used by the researcher to be able to calibrate the motion tracking software in order to approximate the total distance the surgical instrument has travelled (Figure 1A). In a few video fragments, the Intuitive Surgical Large Needle Driver was used by the right robotic arm. In this case, calibration was done by the researcher using this instrument's jaw length, consisting of 1 cm (Figure 1B).

After calibration the researcher only had to manually place a tracking point on the instrument's first joint and press play, the Kinovea software was designed to be able to follow this tracking point during the procedure based on the selected pixels. The entire surgical procedure from the opening of the peritoneum to the completion of the bladder-urethra anastomosis was automatically tracked by the software (Figure 1A,C). Frequently, the software was not able to follow the selected pixels correctly, which was then manually corrected by the researcher by moving the tracking point back to the originally selected point on the instrument. The instrument tracking was manually interrupted if there was something covering it, such as tissue, blood or another instrument. It was also stopped if the instrument was out of view of the camera or if the camera was in movement. If a pelvic lymph node dissection was performed, the instrument was not tracked during this part of the surgery, since not all videos contained a pelvic lymph node dissection.

2.7 | Patient selection and matching

Patients were selected based on their pre-operative and post-operative continence as defined based on the International Consultation Incontinence Modular Questionnaire -Short Form (ICIQ-SF score). The ICIQ-SF is a PROMs questionnaire which registers the patient's urinary incontinence on three domains, the frequency of urine leakage (0-5 points), the amount of leakage according to the
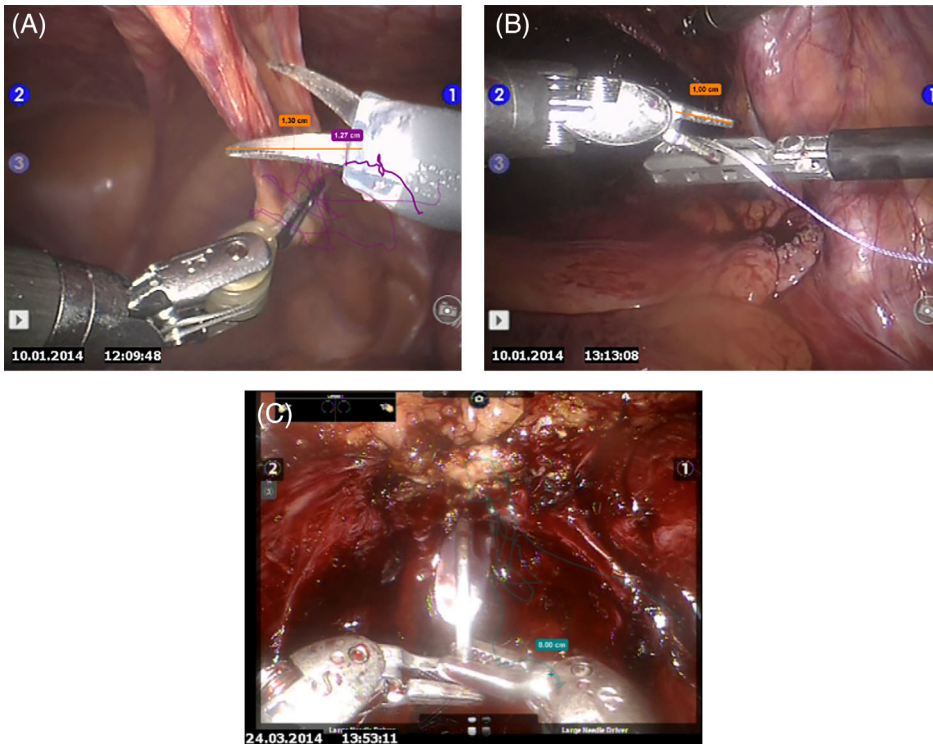
FIGURE 1 A) Example of a calibration line and tracking point on the Monopolar Curved Scissors. B) Example of a calibration line on the Large Needle Driver. C) Example of a tracking point on the Large Needle Driver
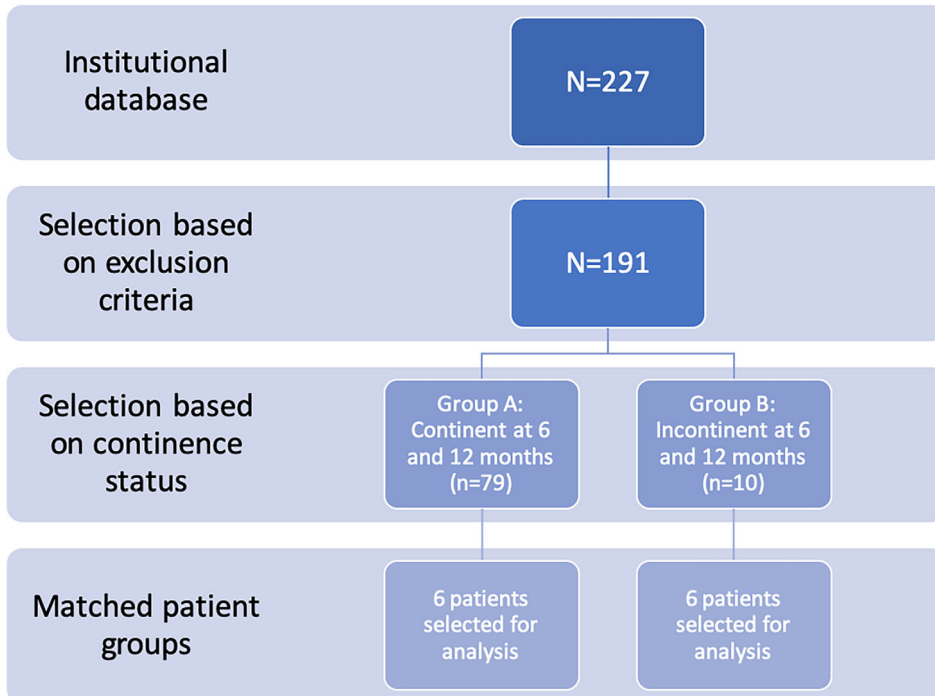

FIGURE 2 Study flowchart

patient (0-6 points) and the interference of the urine leakage with everyday life (0-10 points). An additional question that asks in which situation the urine leaks gives more insight into the type of incontinence. The cumulative scores of the three questions (0-21 points) represent the patients' experience of urinary incontinence. In this study, an ICIQ-SF score of 0 at 6 and 12 months post-operative was defined as continent, while and ICIQ-SF of >10 at 6 and 12 months post-operative was defined as incontinent. Two groups of patients were selected. The first group (A) consisted of patients who were continent pre-operative and were continent at 6 and 12 months after surgery this group was matched with a second group (B) of patients who were continent pre-operative and were incontinent at 6 and 12 months after surgery.
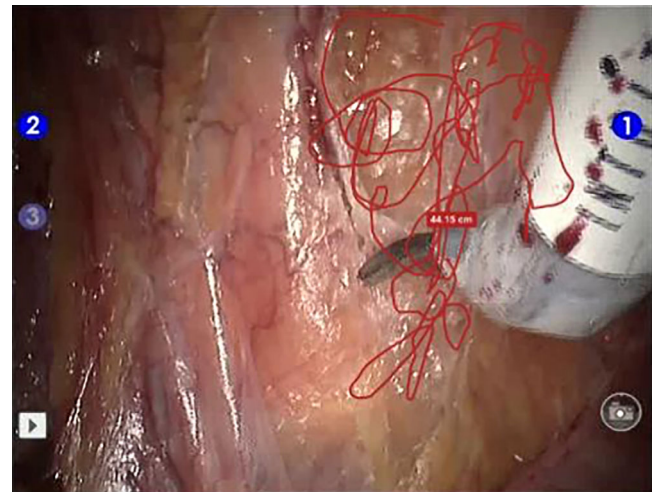
TABLE 2 Baseline characteristics of the patients

Characteristics	Post-operative incontinent patients (n = 6) Median (min-max)	Post-operative continent patients (n = 6) Median (min-max)	P-value	Z-value
Age (years)	64 (57-67)	62.5 (53-66)	0.674	−0.420
Body mass index (kg/m^2)	26.86 (20.45-32.55)	25.99 (22.50-32.11)	0.917	−0.105
Prostate size (ml)	50 (39-81)	38 (35-82)	0.225	−1.214
Gleason score	7 (6–8)	6 (6-7)	0.157	−1.414
Membranous urethral length (mm)	11.07 (8.69-13.10)	13.58 (10.48-16.15)	0.273	−1.095
Surgery date, median (Inter Quartile Range)	15 February 2014 (2 February 2013 to 15 March 2015)	27 April 2014 (12 June 2012 to 29 January 2015)	0.600	−0.524
Pre-operative IPSS score	3 (0-7)	2.5 (0-8.5)	0.892	−0.135
Six-month post-operative IPSS score	15 (0-19)	3 (0-5)	0.042	−2.032
Twelve-month post-operative IPSS score	10 (6-16)	2 (0-5)	0.043	−2.023
Pre-operative ICIQ score	0 (0-0)	0 (0-0)	1.000	0.000
Six-month ICIQ score	15 (11-20)	0 (0–0)	0.027	−2.207
Twelve-month ICIQ score	14 (12-17)	0 (0-0)	0.026	−2.232

Abbreviations: ICIQ, International Consultation Incontinence Modular Questionnaire -Short Form; IPSS, International Prostate Symptom Score.

TABLE 3 Total time analysed, duration of the surgery, percentage of surgery analysed, automatic and calculated total path length in cm, difference in total path length in %, number of sudden movements and average speed

	Median	Min	Max
Total time analysed (minutes)	18.22	8.13	27.05
Duration of the surgery (minutes)	81.00	57.00	99.00
Percentage of surgery analysed (%)	21.74	14.26	29.73

FIGURE 3 Example of a fully tracked trajectory of the Monopolar Curved Scissors

The patients in the continency group were matched according to the date of the surgery, the age of the patient, BMI of the patient and the pre-operative intentions of saving the neurovascular bundles during surgery. All incontinent patients were manually compared to the continent patients by two individual researchers (A.J.W.B. and H.F.N.). Based on the number of variables in which the pairs matched a matching score of zero to four was given to the patients, each matched variable resulted in a point in the total matching score. The patients were matched based on age (difference of ≤5 years = 1 matching point), body mass index (BMI) (difference ≤ 3 points = 1 matching point), date of the surgery (difference ≤ 3 months = 1 matching point) and pre-operative intentions of saving the neurovascular bundles during surgery on both sides (Neuro Vascular Bundle (NVB) sparing the same in both patients = 1 matching point). A matching score of four was the best possible match. Based on the matching scores the best matched patient pairs were selected for analysis, since almost no perfect matches existed (Appendix A). If matched pairs with similar matching scores existed, a definitive choice was made based on the variable on which the patients matched (Appendix A).

2.8 | Power analysis

Ganni et al[8] observed both experts and novice participants during a basic laparoscopic cholecystectomy procedure using semi-automated video motion tracking via the Kinovea system in order to determine their path length, average instrument movement and number of sudden or extreme movements. This study shows a lower path length for the experts compared to the novices. These results show it is possible to detect a difference in populations' means of 60 cm (127 cm for experts, 187 cm in novice analysis) in total path length. For this study, we assume the path length in the incontinent patients is similar to that of a novice and the path length in the continent patients is similar to that of an
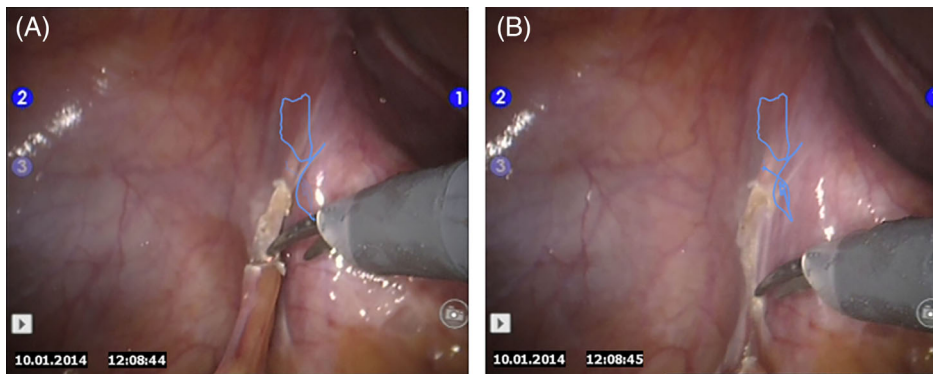
FIGURE 4 Tracking point on instrument (A) and tracking point on tissue 48 milliseconds later (B)

TABLE 4 Automatic and calculated total path length in centimetre, difference in total path length in %, number of sudden movements and average speed

	Median	Min	Max
Automatic total path length (cm)	1594.95	949.91	3264.24
Calculated total path length (cm)	2134.99	1585.38	3515.76
Median difference in path length (%)	16.10		
Automatic number of sudden movements	0	0	0
Calculated number of sudden movements	101	45	167
Automatic average speed (cm/s)	1.67	1.22	2.05
Calculated average speed (cm/s)	2.16	1.48	2.63
Median difference in average speed (%)	23.17		

expert. Based on a power calculation using .05 as alpha, a power of 0.80, and an effect size of 60 cm a sample size of four patients per subgroup would be sufficient for the main objective of this study.

2.9 | Data analysis

Data analysis was performed using SPSS statistics v24 (IBM, NY). Frequency statistics were calculated for patient demographic data and surgeon scores. Correlation between Kinovea results on the one hand and ICIQ-scores, IPSS scores and EORTC QLQ-PR25 scores were calculated using a Spearman's Rho test. The Wilcoxon signed-rank test was used to compare differences in results between the matched patient groups. The McNemar's test was used in order to compare differences in results between the matched patient groups in case of dichotomous variables. Statistical significance was set at $P < .05$ based on a two-tailed comparison.

3 | RESULTS

After exclusion based on the exclusion criteria 191 of the 227 patients were eligible for inclusion (Figure 2). Based on continence status 79 patients were included in group A, 10 patients in group B.

After selection, a total of six incontinent and six continent patients were selected and matched based on age, BMI, prostate size, Gleason score, clinical and pathologic tumour stage, membranous urethral length, IPSS, ICIQ and EORTC QLQ-PR25 (Table 2). There were no significant differences between the incontinent and continent group in baseline characteristics, including age, BMI, date of surgery, oncological data and membranous urethral length. The pre-operative PROMs, that is, ICIQ-SF, IPSS and EORTC QLQ-PR25 also showed no difference between the incontinent and continent group.

3.1 | Results of automated surgical movements assessment using Kinovea

Every surgical video was tracked from the opening of the peritoneum to the completion of the bladder-urethra anastomosis. Of the median duration of the surgery (81.00 minutes), only a median length of 18.22 minutes of the video could be tracked (median 21.74%) (Table 3). Figure 3 shows an example of a trajectory of the movements of the Monopolar Curved Scissors tracked using Kinovea.

We found the computer program could not run fully automatic in RARP surgery since the tracking point could often not correctly identify the instrument's joint, meaning the tracking point had to be manually placed and moved by the researcher. To illustrate this, in Figure 4, a tracking point was placed on the instrument manually (Figure 4A) after which the tracking program ran automatically. Forty-eight milliseconds later, the tracking point had moved to a pixel in the surrounding tissue (Figure 4B).

Table 4 shows the automatic and manually calculated total path length, sudden instrument movements and average speed. The total path length and average speed manually calculated using formulas from Ganni et al's article[8] differed from the automatically calculated results. The median difference in total path length was 16.10% and the median difference in average speed was 23.17%. This data has been summarized in box-and-whisker plots in Figure 5A and Figure 5B

3.2 | Relation between Kinovea results and post-operative continence status

When comparing the 6- and 12-month post-operative IPSS and 6 months post-operative EORTC QLQ-PR25 score the post-operative

FIGURE 5 Summary data of the automatic vs calculated total path length in cm (A) and the automatic vs calculated average speed in cm/s (B)
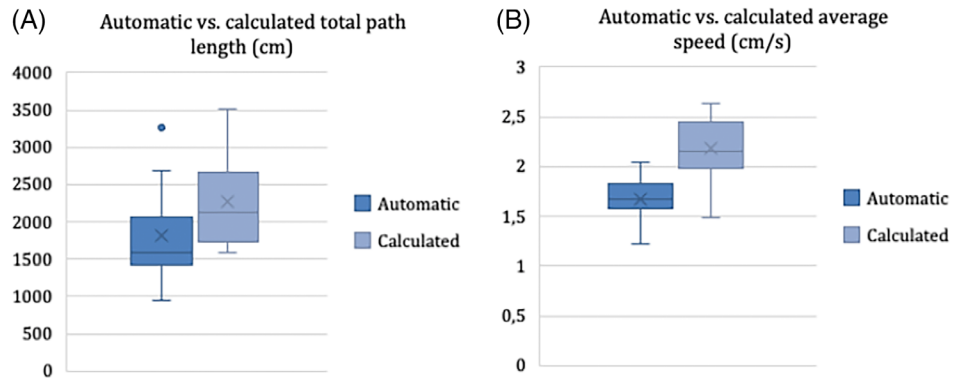

(A) Automatic vs. calculated total path length (cm)

(B) Automatic vs. calculated average speed (cm/s)

TABLE 5 Results of patient reported outcome measures (PROMs) at 6 and 12 months post-operative in arm 1, incontinent vs continent

	Post-operative incontinent patients (n = 6) Median (min-max)	Post-operative continent patients (n = 6) Median (min-max)	P-value	Z-Value
Six-month post-operative IPSS score	15 (0–19)	3 (0-5)	0.042	−2.032
Twelve-month post-operative IPSS score	10 (6–16)	2 (0-5)	0.043	−2.023
Six-month ICIQ score	15 (11-20)	0 (0-0)	0.027	−2.027
Twelve-month ICIQ score	14 (12-17)	0 (0-0)	0.026	−2.232

Abbreviations: ICIQ, International Consultation Incontinence Modular Questionnaire -Short Form; IPSS, International Prostate Symptom Score.

TABLE 6 Correlations between Kinovea measurements and post-operative patient reported outcome measures (PROMs) at 6 and 12 months post-operative

	Total path length (cm)		Average speed (cm/s)	
	Correlation	P-value	Correlation	P-value
Six-month post-operative IPSS score	0.035	0.913	0.311	0.324
Twelve-month post-operative IPSS score	0.071	0.845	0.310	0.383
Six-months ICIQ score	0.092	0.766	0.193	0.547
Twelve-month ICIQ score	0.155	0.631	0.190	0.555

Abbreviations: ICIQ, International Consultation Incontinence Modular Questionnaire -Short Form; IPSS, International Prostate Symptom Score.

continent group had significantly lower IPSS scores and lower EORTC QLQ-PR25 scores at 6 and 12 months after surgery (Table 5). Based on the selection of patients the ICIQ score at 6 and 12 months after surgery also showed a significant difference.

When correlating the results of the manual Kinovea calculations with the post-operative PROMs (ie, the IPSS scores, EORTC QLQ-PR25 scores and ICIQ scores) no significant correlations were found (Table 6).

3.3 | Relation between Kinovea results of different phases of the surgery and post-operative continence status

Of the seven surgical steps defined in the PROTEST assessment method[6] the median speed was highest during the pelvic floor muscle exposure on the left side (2.8 cm/s) and lowest during nerve preservation on the left side (1.51 cm/s) (Table 7). The median percentage analysed was the highest during the bladder neck dissection (41.14%) and the lowest during the ureterovesical anastomosis (6.86%).[6]

The Kinovea results of the "Pelvic floor muscle exposure/opening of the endopelvic fascia" (Table 8), and the "urethrovesical anastomosis" (Table 9) phases to post-operative PROMs showed no significant correlations.

4 | DISCUSSION

The aim of this study was to determine whether Kinovea is a valid tool for automated surgical movements tracking in RARP and if results could be used to evaluate a possible relation between surgical movements and post-operative urinary continence after RARP.

4.1 | The value of automated assessment of surgical movements in RARP surgical videos using Kinovea

To answer whether Kinovea is a valid tool for assessment of surgical movements in RARP surgical videos, an analysis of 12 entire RARP

TABLE 7 Kinovea metrics for each of the seven surgical steps defined in the PROTEST assessment method[5]

RARP step	Number of videos in which Kinovea data was available	Median speed (cm/s) (min-max)	Median time analysed (seconds) (min-max)	Median phase duration (seconds) (min-max)	Median percentage of phase analysed (%) (min-max)
Pelvic floor muscle exposure RIGHT	11	2.61 (1.69-3.54)	9.12 (4.52-80.96)	63 (35-202)	18.38 (4.04-72.59)
Pelvic floor muscle exposure LEFT	7	2.88 (1.93-3.28)	18.56 (3.56-44.64)	90 (25-138)	26.11 (5.93-51.31)
Pelvic floor muscle exposure COMBINED	7	2.68 (1.88-3.41)	36.40 (12.12-125.6)	146 (60-289)	24.53 (9.23-43.46)
Bladder neck dissection	12	1.89 (1.46-2.28)	207.78 (99.44-382.40)	447 (286-1044)	41.14 (23.49-62.06)
Ligation of prostatic pedicles RIGHT	12	2.07 (1.38-3.28)	28.48 (18.84-73.12)	188 (74-435)	16.60 (9.88-38.89)
Ligation of prostatic pedicles LEFT	11	1.71 (0.95-3.25)	21.92 (5.92-50.40)	143 (53-370)	15.77 (4.14-31.27)
Ligation of prostatic pedicles COMBINED	11	1.81 (1.26-3.00)	59.84 (26.00-93.36)	281 (222-805)	15.60 (7.67-29.99)
Nerve preservation RIGHT	9	2.18 (1.17-2.47)	19.20 (4.04-72.16)	113 (46-600)	12.03 (5.38-61.48)
Nerve preservation LEFT	7	1.51 (1.42-1.93)	20.92 (2.88-45.56)	190 (89-603)	8.72 (3.24-21.90)
Nerve preservation COMBINED	6	1.84 (1.29-1.96)	52.34 (15.60-96.48)	337 (228-764)	11.82 (6.61-16.68)
Management of prostatic apex/ urethra	12	1.83 (1.17-2.79)	125.62 (14.52-281.64)	438 (211-1070)	28.63 (6.88-38.87)
Urethrovesical anastomosis	12	1.75 (0.98-2.29)	38.58 (7.80-187.28)	648 (386-1093)	6.86 (1.40-19.27)

TABLE 8 Correlations between Kinovea measurements of the "pelvic floor muscle exposure/opening of the endopelvic fascia" and post-operative patient reported outcome measures (PROMs) at 6 and 12 months post-operative

	Right side of the patient				Left side of the patient			
	Total path length (cm)		Average speed (cm/s)		Total path length (cm)		Average speed (cm/s)	
	Correlation	P-value	Correlation	P-value	Correlation	P-value	Correlation	P-value
Six-month post-operative IPSS score	0.264	0.432	0.332	0.319	−0.504	0.249	0.502	0.251
Twelve-month post-operative IPSS score	0.441	0.202	−0.072	0.844	−0.495	0.318	0.055	0.917
Six-month ICIQ score	0.049	0.886	0.023	0.946	−0.646	0.117	0.074	0.875
Twelve-month ICIQ score	0.034	0.921	0.007	0.984	−0.642	0.120	0.040	0.933

Abbreviations: ICIQ, International Consultation Incontinence Modular Questionnaire -Short Form; IPSS, International Prostate Symptom Score.

procedures was performed by a trained researcher. The analysis itself was not technically challenging, after manual calibration, the researcher only had to position the tracking point on the most distal hinge of the instrument and press play.

During automated analysis, the videos had to be manually adjusted since Kinovea was frequently not able to run fully automatic. In most instances the short distance between camera and tissue (level of zoom), the speed of the surgical instruments, tissue overlapping the

TABLE 9 Correlations between Kinovea measurements of the "urethrovesical anastomosis" and post-operative patient reported outcome measures (PROMs) at 6 and 12 months post-operative

	Total path length (cm)		Average speed (cm/s)	
	Correlation	P-value	Correlation	P-value
Six-month post-operative IPSS score	−0.196	0.541	0.425	0.169
Twelve-month post-operative IPSS score	−0.328	0.355	0.529	0.116
Six-month ICIQ score	−0.318	0.313	0.006	0.985
Twelve-month ICIQ score	−0.331	0.293	0.141	0.661

instrument and the fact that the instrument has flexible joints, made it difficult to track the instrument correctly. These findings show that it is difficult to use Kinovea for assessment of surgical movements in RARP surgical videos.

The group of Ganni et al was able to automatically track the instrument during surgery with Kinovea. They checked a few minutes manually per video in order to confirm their results.[8] We attempted to perform similar analysis in robot-assisted surgery as in the study of Ganni et al. The analysis was based on the same principal by relating the distance moved by the surgical instrument to the relative measurements of the instrument. In our study the frequent manual adjustments raised the need for continued monitoring by the researcher, this meant it was a very time-consuming method of analysis.

In this study, we compared two forms of output from the Kinovea program, one form is the automated results (automated output) given by the Kinovea program and the other form of output was the calculated results based on the raw Kinovea data using the formulas as described in the study by Ganni et al (manual output). An interesting finding was the lack of sudden movements reported in the videos when using the automated output, that did not match the number of sudden movements found using the manual calculations based on the formulas used by Ganni et al.[8] The automated and manual total path length and average speed also did not match.

These inconsistencies raise the questions why the outputs differ and which method is more reliable. When checking these manual sudden movements, they appeared to be caused by manual repositioning of the tracking point on the instrument by the researcher after the tracking point had lost the instrument's joint during automated analysis by the Kinovea program. During the assessment of the videos, an average of 101 manual repositioning's were registered. This makes the sudden movements unreliable for analysis in robot-assisted surgery, which could mean the automated results represent a more accurate analysis of the surgical movements. The group of Ganni et al did not report this problem, either because no manual repositioning was necessary in their study or because they did not manually review the entire surgical video.[8]

4.2 | Relation between Kinovea results and post-operative continence status

This study shows there is no relation between the motion tracking results and the continence status of this patient. The main reason could be the limited quality of the Kinovea motion tracking results since on average less than a quarter of the videos could be tracked. Another explanation could be there is no relation between the total path length, average speed of the instrument and the continence status of the patient.

4.3 | Relation between Kinovea results of different phases of the surgery and post-operative continence status

The correlation between Kinovea results of the "pelvic floor muscle exposure/opening of the endopelvic fascia" and the "urethrovesical anastomosis" to the PROMs results also did not give any significant correlations.

Since the value of the results using the Kinovea program is questionable, the lack of correlation could be due to the fact that only limited parts of the surgery were analysed or it could mean the results obtained using Kinovea cannot be used to predict post-operative continence. Since the Kinovea software was not able to automatically assess surgical movements in RARP it proved to be an invalid tool for automated surgical movement tracking in robot-assisted surgeries.

4.3.1 | Strengths and limitations

To the best of our knowledge, this is the first study into the possibilities of using existing surgical videos of real RARP procedures to identify the effects of surgical movements on post-operative outcome. This study shows the results of the first use of Kinovea as a software-based surgical motion tracking tool in robot-assisted surgery. Although the assessment of the movements using Kinovea was time-consuming and had its challenges, Kinovea does give the researcher the means to simultaneously assess the anatomy during analysis of the surgical videos. One of the alternatives to Kinovea, the dVlogger, only provides raw movement data without the ability to correlate this to the surgical videos.[9]

The data represented in this study are the results of the parts were automated tracking was possible, no data was interpolated. During tracking the software was frequently not able to follow the selected pixels correctly, which was then manually corrected by the researcher by moving the tracking point back to the originally selected point on the instrument. The movement data due to the repositioning of the tracker point was deleted from the data after manual verification and checking of the path in the video file.

The place of the surgery within the learning curve of the surgeon could influence the results of this type of analysis. Other studies have shown that arm movement analysis can be used to separate beginning surgeons from experts.[8,15] In this study, the selection of cases has been adjusted to take into account when the surgery was performed in order to reduce the influence of learning curve on the results of this study.

The Kinovea program was able to assess surgeons' skills using existing surgical videos rather than in a simulator, without needing extra equipment for movement tracking in laparoscopic surgery.[8] The analysis in robot-assisted surgery does not appear to be as valid as the analysis in laparoscopic surgery. In robot-assisted surgery the frequent manual replacement and moving of the tracking point during the Kinovea analysis adds a subjective component to an otherwise objective measurement.

Further research with larger groups of patients and a different automated tracking system is needed in order to investigate the relation between surgical movements, surgical skills and post-operative outcomes. A combination of video assessment and dVlogger data could hold the key to find metrics related to post-operative outcome. To the knowledge of this group no such analysis method is currently available to be used in a retrospective analysis without additional equipment for tracking the movements. The use of artificial intelligence in combining both video assessment and surgical movement assessment could eliminate human interference and lead to an objective and automated assessment of the surgical video.

5 | CONCLUSION

Kinovea can be used to retrospectively assess instrument movement in laparoscopic surgery without requiring extra equipment.[8] In this study, Kinovea was used to assess if this software could be used to automatically measure surgical movement in RARP videos. Based on the results of this study, because of a more close-up camera position for robotic as compared to laparoscopic surgery, the speed of the surgical instruments, tissue overlapping the instrument and the fact that the instrument has flexible joints, the Kinovea software cannot be used to automatically assess surgical movements in RARP surgical videos.

ORCID

Alexander J. W. Beulens https://orcid.org/0000-0002-7105-1011

REFERENCES

1. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997;84(2): 273-278.
2. Birkmeyer JD, Finks JF, O'reilly A, et al. Surgical skill and complication rates after bariatric surgery for the Michigan bariatric surgery collaborative. *N Engl J Med*. 2013;369:1434-1442.
3. 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.
4. Hussein AA, Ghani KR, Peabody J, et al. Development and validation of an objective scoring tool for robot-assisted radical prostatectomy: prostatectomy assessment and competency evaluation. *J Urol*. 2017; 197(5):1237-1244.
5. Husslein H, Shirreff L, Shore EM, Lefebvre GG, Grantcharov TP. The generic error rating tool: a novel approach to assessment of performance and surgical education in gynecologic laparoscopy. *J Surg Educ*. 2015;72(6):1259-1265.
6. Beulens AJW, Brinkman WM, Van der Poel HG, et al. Linking surgical skills to post-operative outcomes: a Delphi study on the robot-assisted radical prostatectomy. *J Robot Surg*. 2019;13:675-687.
7. Goldenberg MG, Goldenberg L, Grantcharov TP. Surgeon performance predicts early continence after robot-assisted radical prostatectomy. *J Endourol*. 2017;31(9):858-863.
8. Ganni S, Botden SMBI, Chmarra M, Goossens RHM, Jakimowicz JJ. A software-based tool for video motion tracking in the surgical skills assessment landscape. *Surg Endosc*. 2018;32(6):2994-2999.
9. Hung AJ, Chen J, Jarc A, Hatcher D, Djaladat H, Gill IS. Development and validation of objective performance metrics for robot-assisted radical prostatectomy—a pilot study. *J Urol*. 2017;199(1):296-304.
10. Hung AJ, Chen J, Che Z, et al. Utilizing machine learning and automated performance metrics to evaluate robot-assisted radical prostatectomy performance and predict outcomes. *J Endourol*. 2018;32(5):438-444.
11. http://www.urospec.com/uro/Forms/ipss.pdf.
12. van Andel G, Bottomley A, Fosså SD, et al. An international field study of the EORTC QLQ-PR25: a questionnaire for assessing the health-related quality of life of patients with prostate cancer. *Eur J Cancer*. 2008;44(16):2418-2424.
13. https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf.
14. Chang YJ, Chang CH, Peng CL, et al. Measurement equivalence and feasibility of the EORTC QLQ-PR25: paper-and-pencil versus touch-screen administration. *Health Qual Life Outcomes*. 2014;12(1):1-10.
15. Hung AJ, Oh PJ, Chen J, et al. Experts vs super-experts: differences in automated performance metrics and clinical outcomes for robot-assisted radical prostatectomy. *BJU Int*. 2018;123(5):861-868.

APPENDIX A | MATCHING SCORES PER PAIR IN THE CONTINENCY GROUP

The patients were matched based on age (difference of ≤5 years = 1 matching point), BMI (difference ≤3 points = 1 matching point), date of the surgery (difference ≤90 days = 1 matching point) and pre-operative intentions of saving the neurovascular bundles during surgery on both sides (NVB sparing the same in bot patients = 1 matching point).

A total 12 possible match pairs were identified for the 10 incontinent patients. The match pairs 82/84, 128/129 and 212/213 were chosen because of their maximum match score of four. During review the video of patient 129 did not work, this was a reason to exclude this matched pair and replace it with another matched pair. Individual patients were part of multiple matched pairs in case of five matched pairs (147/155, 140/155, 140/143,

116/121 and 116/107). Of these matched pairs 140/155 was not chosen because match pair 140/143 matched better on date of the surgery, which reduces the influence of the learning curve on the post-operative results. Match pair 147/155 was chosen because there is a shorter interval between dates of the surgery compared to 140/155. Match pair 116/121 was not chosen because match pair 116/107 matched similar on date of the sur-

gery and better on BMI of the patient, which could reduce the influence of BMI on the post-operative results. The match pairs 140/143, 167/153 and 116/107 were chosen because they matched on the date of the surgery and BMI. Match pair 200/178 was not chosen because of the large interval between dates of the surgery, which could increase the influence of the learning curve on the post-operative results.

Study # incontinent	Study # continent pair	Difference OR date (days)	Match point Operation date	Difference BMI (points)	Match point BMI	Difference age (years)	Match point age	Difference NVB	Match point NVB	Total score match point	Matched pair selected (yes/no)
82	84	38	1	1.20	1	3	1	None	1	4	Y
212	213	1	1	2.05	1	3	1	None	1	4	Y
140	143	42	1	1.45	1	13	0	None	1	3	Y
167	153	89	1	0.59	1	4	1	None	1	3	Y
116	107	90	1	0.44	1	7	0	None	1	3	Y
147	155	130	0	1.42	1	1	1	None	1	3	Y
128	129	7	1	1.23	1	1	1	None	1	4	N
136	141	50	1	0.15	1	10	0	None	1	3	N
140	155	203	0	0.88	1	7	0	None	1	2	N
32	30	2	1	0.13	1	13	0	None	1	3	N
200	178	276	0	0.02	1	2	1	None	1	3	N
116	121	49	1	5.13	0	0	1	None	1	3	N