Psychomotor skills assessment by motion analysis in minimally invasive surgery on an animal organ

Erlend Fagertun Hofstad (MSc) [1,2], Cecilie Våpenstad (MSc) [1,2,3], Lars Eirik Bø (MSc) [1,2,3], Thomas Langø (PhD) [1,2], Esther Kuhry (PhD) [4,5], Ronald Mårvik (PhD) [2,4,5,6]

[1]*Department of Medical Technology, SINTEF Technology and Society, Trondheim, Norway*

[2]*St. Olavs Hospital, Trondheim University Hospital, The Norwegian National Advisory Unit for Advanced Laparoscopic Surgery, Trondheim, Norway*

[3] *Department of Circulation and Medical Imaging, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

[4]*Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

[5]*Department of Surgery, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway*

[6]*National Center for Advanced Laparoscopic Surgery, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway*

Corresponding author:
Erlend Fagertun Hofstad, MSc, Research Scientist
SINTEF Technology and Society
Dept. of Medical Technology
PB 4760 Sluppen
7465 Trondheim, Norway
Phone: +47 911 96 173
e-mail: erlend.hofstad@sintef.no

Psychomotor skills assessment by motion analysis in minimally invasive surgery on an animal organ

BACKGROUND: A high level of psychomotor skills is required to perform minimally invasive surgery (MIS) safely. To be able to measure these skills is important in the assessment of surgeons, and as it enables constructive feedback during training. The aim of this study was to test the validity of an objective and automatic assessment method using motion analysis during a laparoscopic procedure on an animal organ.

METHODS: Experienced surgeons in laparoscopy (experts) and medical students (novices) performed a cholecystectomy on a porcine liver box model. The motions of the surgical tools were acquired and analyzed by 11 different motion-related metrics, i.e. a total of 19 metrics as 8 of them were measured separately for each hand. We identified which of the metrics the experts outperformed the novices.

RESULTS: In total two experts and 28 novices were included. The experts achieved significantly better results for 13 of the 19 instrument motion metrics.

CONCLUSIONS: Expert performance is characterized by a low *time* to complete the cholecystectomy, high *bimanual dexterity* (instrument coordination), a limited amount of movement and low measurement of *motion smoothness* of the dissection instrument, and relatively high usage of the grasper to optimize tissue positioning for dissection.

Introduction

Psychomotor skills are essential when performing minimally invasive surgery (MIS) [1-3]. To ensure that surgeons have the required level of psychomotor skills to perform surgery safely, assessment of these skills is required [4]. Assessment of skills is also valuable to provide feedback during training [5] and to verify the trainee's development during a training program [6]. In addition, objective and automatic measurement of the psychomotor skills gives a fair and reliable assessment with limited effort from the trainer. Tracking the position of the surgical instruments can be used to calculate motion-related metrics such as *path length* and *motion smoothness* [7-9]. By looking at performances of surgeons with different level of experience, metrics that measure relevant psychomotor skills for MIS can be identified [7, 10]. Monitoring and analyzing motions of surgical instruments in the operating room can, in the future, potentially be used to indicate adverse, harmful events ahead of their occurrence.

Our goal with this study was to develop a valid, objective and automatic assessment method that is capable of measuring surgical performance during laparoscopic procedures using real tissue organs. The assessment is based on motion analysis of the surgical instruments that were used to perform a cholecystectomy on a porcine liver placed in a box trainer. We tested a set of motion-related metrics for their validity to assess psychomotor skills in MIS. We have already tested most of the metrics in a previous study with a different setup [7], and several of the metrics have also been presented in other papers [9]. However, to our knowledge, the calculation of the total path length of one instrument relative to the total path length of both instruments has previously not been published. There is also novelty in the investigation of the relevance of these metrics in a setup with real tissue organs.

Significant effort has been invested in identifying instrument motion metrics that are valid in terms of measuring the performance on virtual reality and analog simulators [9]. However, limited research has been performed regarding automatic assessment of psychomotor skills while operating on real tissue. Smith et al. [11] studied motion analysis of the surgical instruments during cholecystectomies in liver models in the laboratory, performed by surgeons with different level of experience. They measured the performance by the metrics *time*, *path length*, *velocity* and *number of movements*. Aggerwal et el. [12] measured the same metrics during cholecystectomies on patients. In this study, we measured the performance of experienced laparoscopic surgeons (experts) and medical students (novices) by using the four metrics tested in the studies by Aggerwal [12] and Smith [11], in addition to another seven metrics.

It is essential to validate all setups used for training and assessment, to ensure that the desired set of skills are acquired and tested. Construct validity is evaluating if a testing instrument is able to identify the constructs, i.e. the skills it is intended to measure [10]. We tested construct validity by evaluating whether our setup and assessment metrics were able to differentiate between expert and novice performance.

Materials and Methods

Participants

The participants in the study were expert laparoscopic surgeons (>500 laparoscopic procedures) and novices without any experience in laparoscopic surgery. The latter were recruited from medical students in the 5th and 6th year of medical school (in total six years of study) and interns in their first year of practice. All aspects of the

study were approved by the Norwegian Data Protection Agency, and all subjects gave written informed consent.

Task

The participants' psychomotor skills were measured while they performed a cholecystectomy on a pig liver placed in a box trainer (the Pulsating Organ Perfusion (POP) trainer, Optimist Hg.m.b.H, Innsbruck, Austria), see Fig. 1. The novices were presented a one-hour theoretical introduction to the procedural steps and given a demonstration by an experienced laparoscopic surgeon who performed parts of a cholecystectomy in the training setup. The participants also got the opportunity to become familiar with and test the functionality of the instruments before the procedure. The instruments were: a laparoscopic grasper (Endo Clinch II), an ultrasound (US) hook (Auto Sonix Hook Probe 5 mm) and a clip applier (Endo Clip II ML 10 mm) from Covidien (Mansfield, USA), and scissors (Metzenbaum, Ergo handle, 19 mm jaws) from Olympus GmbH (Hamburg, Germany).

The participants were asked to dissect the Calot's Triangle, identify, clip and cut the cystic duct and artery, and finally dissect the gallbladder from the liver bed. The laparoscopic camera was mounted in a mechanical camera holder and was kept in a fixed position throughout the procedure. As all the participants where right handed, they all held the grasper with their left hand and the US hook, clip applier and the scissors with their right hand.

Motion analysis

Position tracking

The position of the laparoscopic grasper and the US hook were tracked using the Aurora Electromagnetic Measurement System (Northern Digital Inc., Waterloo, ON, Canada). The system consists of a System Control Unit, up to four System Interface units for position sensor inputs and a Field Generator that generates an electromagnetic tracking field.

A position sensor with six degrees of freedom (DOF) was placed close to the tip on the laparoscopic grasper and the US hook, and a reference sensor was placed on the wall of the box trainer, see Fig. 1. The sensors measure positions with an accuracy of 0.48 mm, according to the manufacturer. The positions were sampled at a frequency of 40 Hz.

Due to practical limitations, neither the clip applier nor the cutter was equipped with position sensors. As a consequence, we excluded from the motion analysis all time periods where the US hook was retracted from the box trainer, and replaced with the clip applier or the cutter.
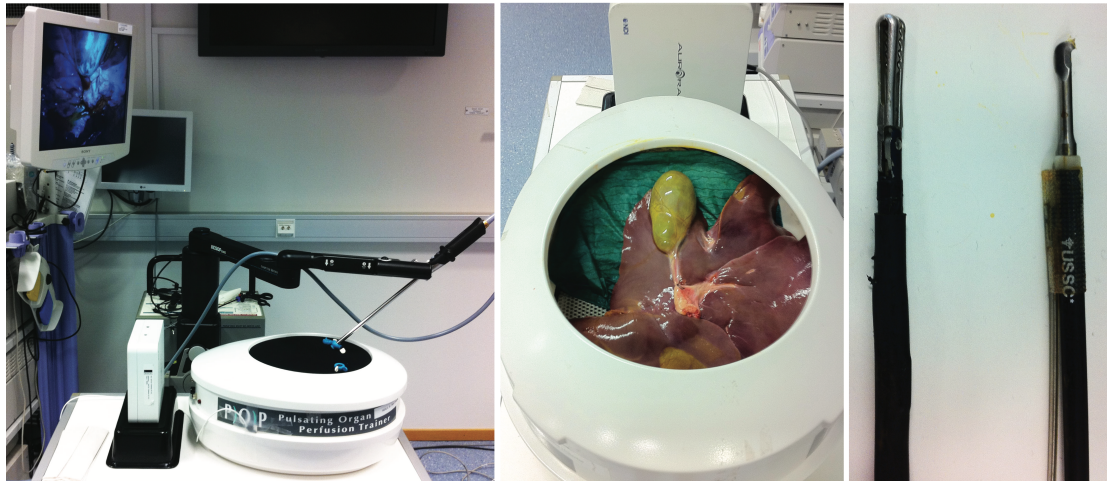
Figure 1. Left: The setup with the box trainer, the laparoscopic camera and the positioning system. Middle: The porcine liver organ placed in the box trainer. Right: The laparoscopic grasper and US hook with a position sensor mounted close to the tip.

Motion analysis metrics

The acquired motion data was analyzed using a previous developed motion analysis software, which was presented in detail in a previous paper [7]. The software was written in MATLAB 7 (The MathWorks Inc., MA, USA).

In order to eliminate the high-frequency background noise, the motion data was filtered using a 6th order low-pass Butterworth filter with cut-off frequency at 6 Hz.

All motion-related metrics were derived from the position and the orientation of the instruments. The position was defined by: $r(t) = [x(t), y(t), z(t)]_{t=0}^{T}$. The orientation was defined by the three angles $[\alpha(t), \beta(t), \gamma(t)]_{t=0}^{T}$, where α and β measure the orientation in two planes perpendicular to the instrument's axis and γ measures the rotation around the instrument's axis, as illustrated in Fig. 2. The following metrics were used in the evaluation of the performance of the participants:

1. *Time* (T): Time from start to completion of the task.

2. *Bimanual dexterity* (BD): The participant's ability to control two instruments at the same time. BD is found by calculating the correlation between the velocity of the tip of the instruments controlled by the left and the right hand:

$$BD = \frac{\int_0^T (v_{left}(t) - \bar{v}_{left})(v_{right}(t) - \bar{v}_{right})dt}{\sqrt{\int_0^T (v_{left}(t) - \bar{v}_{left})^2\, dt \cdot \int_0^T (v_{right}(t) - \bar{v}_{right})^2\, dt}}$$

   where v is the velocity of the instruments and $\bar{v}$ denotes the average velocity over the duration of the exercise.

3. *Path length* (PL): Total movement of the tip of the instrument for the duration of the task, measured in meters:

$$PL = \int_0^T \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt$$

4. *Percentage path length grasper* (PPLG): The path length of the grasper divided by the total path length of both instruments.

5. *Angular length* (AL): The total change in the angle of the tip of the instrument in the plane perpendicular to the instrument's axis, measured in degrees:

$$AL = \int_0^T \sqrt{\left(\frac{d\alpha}{dt}\right)^2 + \left(\frac{d\beta}{dt}\right)^2}\, dt$$

6. *Depth perception* (DP): DP is calculated by the total distance traveled by the tip of the instrument in the instrument's axis direction, measured in meters:

$$DP = \int_0^T \left| \frac{dz}{dt} \right| dt$$

7. *Response orientation* (RO): The total amount of instrument rotation around its axis, measured in degrees:

$$RO = \int_0^T \left| \frac{d\gamma}{dt} \right| dt$$

8. *Motion smoothness* (MS): The total change in acceleration of the tip of the instrument. The motion smoothness is measured in m/s^3, and is normalized by the duration of the task:

$$MS = \sqrt{\frac{1}{2T} \int_0^T \left(\frac{d^3x}{dt^3} \right)^2 + \left(\frac{d^3y}{dt^3} \right)^2 + \left(\frac{d^3z}{dt^3} \right)^2 dt}$$

9. *Number of submovements* (NoS): A submovement is defined by a movement of the tip of the instrument containing a velocity peak of at least 10 mm/s.

10. *Average velocity* (AV): The average velocity of the tip of the instrument for the duration of the task, measured in mm/s.

11. *Idle percentage* (IDLE): Percentage of total time the instrument is moved at a speed < 2 mm/s.
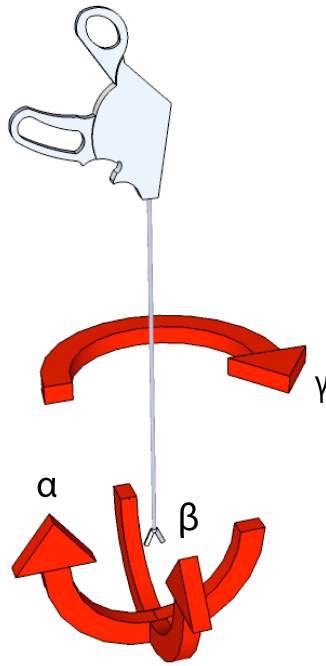
Figure 2. The definition of the orientation of the laparoscopic grasper.

Out of the eleven motion-related metrics, metric number 3 and 5–11 are measured for both hands separately, i.e. a total of 19 metrics were calculated. All metrics but number 4 and 11 were used in the analysis in [7].
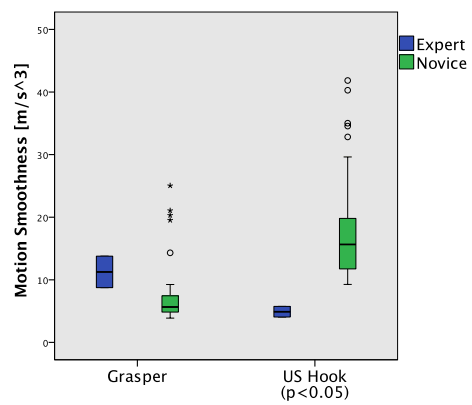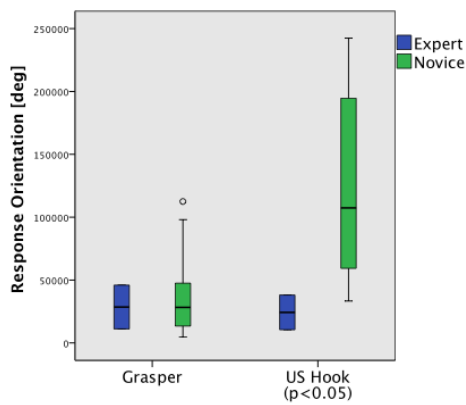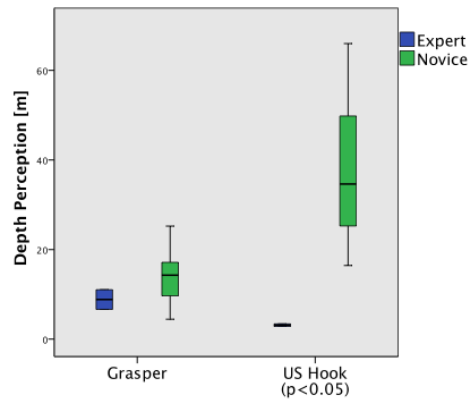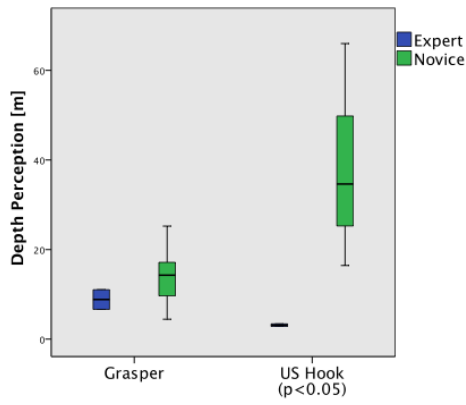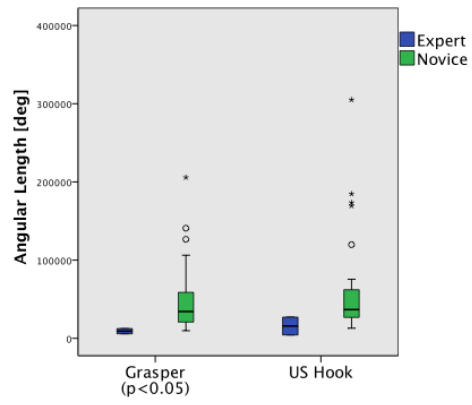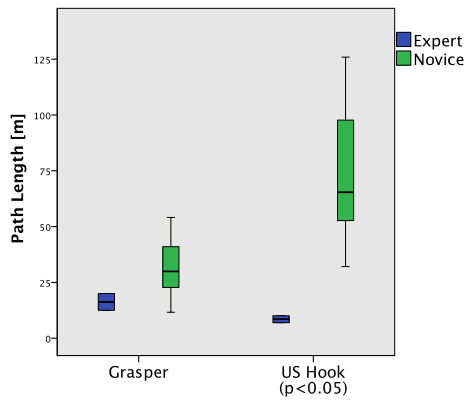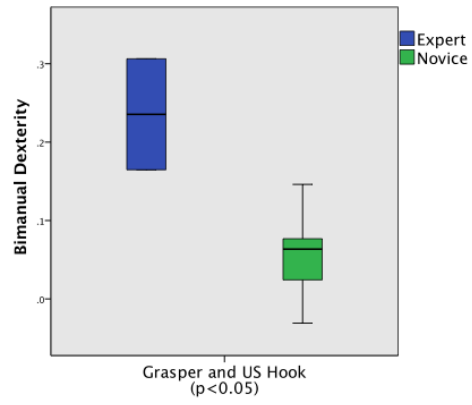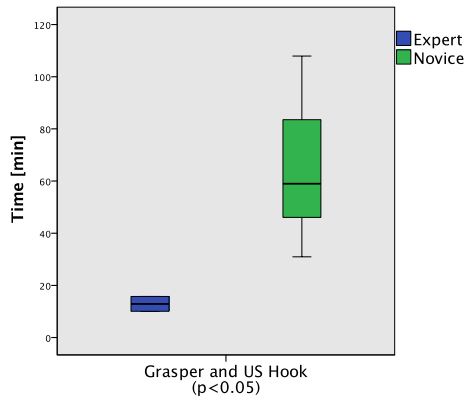
Statistical analysis

The statistical analysis was performed in SPSS v. 23 (IBM Corporation, Armonk, USA). The Mann-Whitney U-test was used to evaluate differences in central tendencies of the two groups: experts and the novices. The Pearson's test was applied to identify correlations between the different metrics.

Results

Two experts and 28 novices were included in the analysis. The results of the motion analysis for the expert and novice groups are presented in Fig. 3 a-k. Metrics with a statistically significant difference between the two groups found by the Mann-Whitney U-test are marked with ($p<0.05$) in Fig. 3 a-k. Statistical significant difference was found for 13 of the 19 metrics tested. Four of the metrics that did not obtain statistical significance were related to the movement of the grasper: *path length*, *depth perception*, *response orientation*, *number of submovements* and *motion smoothness*. The amount of movement of the grasper is thus comparable between the two groups even though the procedural time was about five times longer for the novices.

Regarding the US hook all metrics besides *angular length* showed statistical significance. Also note that the experts moved the grasper at a significantly higher average velocity than the novices. However, they moved the US hook at a significantly lower average velocity.

Correlations between the metrics are presented in Table 1. Several of the metrics are intercorrelated, like *time*, *path length* and *number of submovemnts*. The results from *bimanual dexterity*, *motion smoothness*, *angular length* of the US hook and *idle percentage* of the grasper have less dependency to the other metrics.
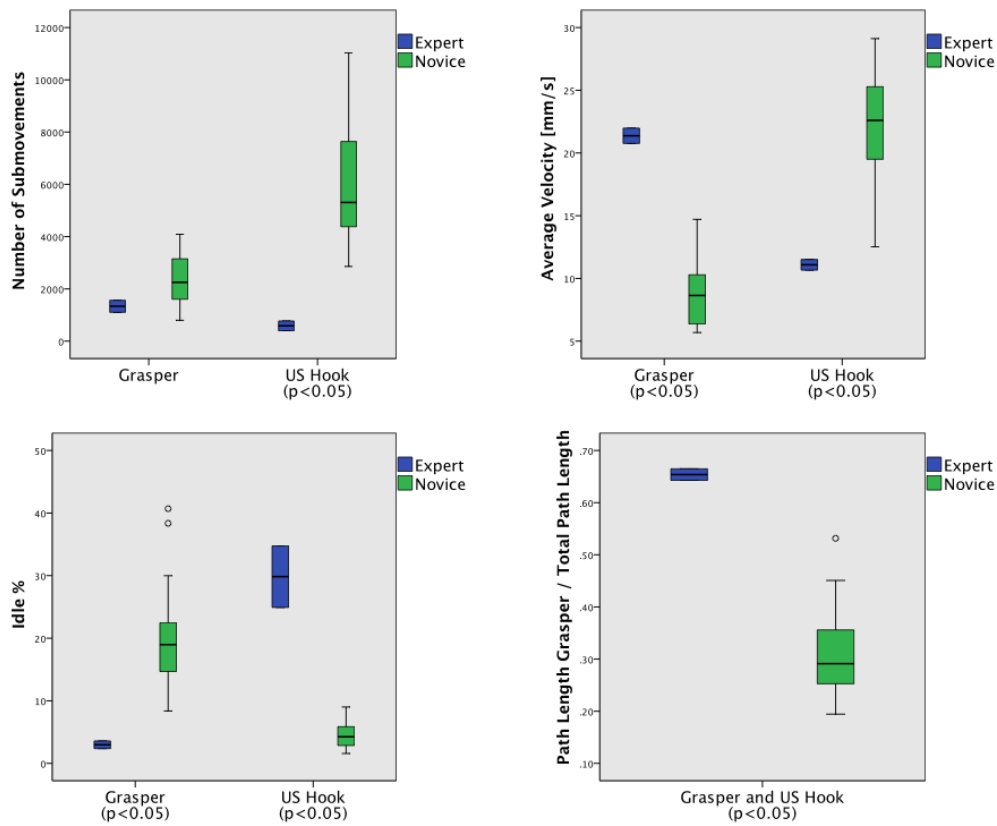
Figure 3 a-k. Motion analysis results from the cholecystectomy task presented as box plots of the metrics. Metrics showing statistically significant difference between the novices and experts are identified with (p<0.05). The middle band shows the median value, the bottom and top of the boxes show the 25th and 75th percentile and the ends of the whiskers show the 5th and 95th percentile. Outliers are plotted as "O" and extreme outliers as "*".

Table 1. The correlation matrix for the motion analysis calculated by Pearson's test. Statistical significant correlations marked with bold numbers.

US Hook	T	BD	PL	AL	DP	RO	MS	NoS	AV
T									
BD	**38**								
PL	**77**	25							
AL	22	0	27						
DP	**73**	21	**99**	29					
RO	**60**	8	**82**	**46**	**83**				
MS	26	30	1	5	2	21			
NoS	**88**	29	**94**	24	**93**	**78**	6		
AV	9	24	**60**	20	**62**	**44**	2	**37**	
IDLE	**49**	**68**	**60**	25	**60**	**45**	22	**61**	**64**

Grasper	T	BD	PL	AL	DP	RO	MS	NoS	AV
T									
BD	**38**								
PL	**71**	24							
AL	**55**	20	**56**						
DP	**68**	20	**96**	**56**					
RO	14	15	**44**	**52**	**47**				
MS	28	2	18	4	9	17			
NoS	**69**	18	**97**	**53**	**95**	**44**	8		
AV	**53**	**57**	5	9	9	22	**51**	7	
IDLE	27	**44**	21	1	26	31	28	29	**67**

Discussion

In this study we investigated the use of motion analysis of laparoscopic instruments to classify performance levels. We compared the performance of two experts in laparoscopic surgery with 28 novices while operating on an animal organ in a box trainer. Eleven different motion analysis metrics were measured. We aimed at identifying metrics for which the experts outperformed the novices. These metrics could thus be used to objectively assess psychomotor skills in the testing setup.

We obtained statistically significant differences in the performances for most of the metrics, with exceptions for *angular length* for the US hook and *path length*, *depth perception*, *response orientation*, *number of submovements* and *motion smoothness* for the grasper. The experts did on average perform the task almost five times faster than the novices, but the average *number of submovements* for the grasper was barely higher for the novices (Fig. 3 a, h). Thus, the experts appear to use the grasper more actively, e.g. to optimize tissue position during the dissection. On the other hand, the experts needed significantly less movement of the US hook to complete the task. Even the *average velocity* of the US hook was lower compared to the novices (about half), which is surprising as the experts completed the dissection about five times faster. This could be explained by the novices' inability to control the instruments, their limited awareness of the patient safety aspect when using a US hook and thus lack of strategy to complete the dissection in an efficient and safe manner. It seemed like the experts used the grasper more actively to position the tissue so that the dissection could continue with minimal movement of the US hook. These findings are partly in contradiction to Smith et el. [11] who also studied motion analysis during cholecystectomies in a pig liver model. They found that both novices and experts made faster and more motions with the right hand, which normally controls the dissection instrument. Another finding by Smith et al. was that novices moved both hands faster than the experts, while in our experiment this was only true for the US hook. The dissimilar results could be explained by the different instrument used to dissect the gall bladder from the liver in the two studies, a US hook in our study and scissors in [11]. The motions were measured differently as well, using sensors on the participant's hands [11] instead of on the instrument tip. Another explanation of the different results might be how the groups are defined. In our study

the novices did not have surgical experience, while in [11] the novice group had performed up to ten cholecystectomies, which might explain the more similar results between experts and novices in [11].

In a similar study by Aggarwal et al. [12], the authors used the ROVIMAS video-based motion tracking device during cholecystectomies on patients, measuring *time*, *path length* and *number of movements*. When including the movements from the entire procedure, they found significant difference between experienced (>100 procedures) and inexperienced (<10 procedures) surgeons for *time* only. The difference in the results from our study can be related to the fact that we separated the movements of the different instruments, and the main difference in the amount of movement was related to one of the instruments in our study, the US hook. Aggarwal et al. [12] included more experts (n=13) and fewer novices (n=6) in the study, and the novices also had more experience, compared to our study. These are aspects that could explain some of the differences in the results.

Bimanual dexterity, which intends to measure the ability to control both instruments at the same time, was significantly higher for the experts than the novices. This supports the assumption that the experts' use of the grasper was more complementary to their use of the US hook. The experts did also perform the task with significantly better *motion smoothness* of the US hook.

In another study performed by our group [7], it was found that good control of an instrument maneuvered by the non-dominant hand and the ability to perform coordinated movements with both hands are not acquired until the surgeons reach a high level of experience. This is supported by the findings in the current study as two of the parameters identified as describing an expert performance best, i.e. *bimanual*

dexterity and *percentage path length grasper*, are related to usage and control of the non-dominant hand.

The correlations between the different motion metrics were calculated and several of the metrics were found to be dependent. It is also logical that e.g. the *number of submovements* is strongly related to the total *path length* of an instrument. However, some of the metrics like *bimanual dexterity* and *motion smoothness* were found to be less correlated to the others, which indicates that they are describing different aspects of dexterity required for minimally invasive surgery. This is an important consideration to be made when finding descriptive motion related metrics that can be used to assess a surgeon's dexterities.

In the study setup the laparoscopic camera was mounted in a fixed position, opposed to a realistic setting where the operator (or assistance) is free to manoeuvre the camera. The setup was chosen in order to be equal to all participants. It would be difficult for the novices to operate the camera themselves, nor was an assistant available to operate the camera. This might have affected the study result, being a disadvantage for the expert surgeons as they are used to optimize the camera position and zoom for different parts of the procedure.

The limited number of experts available to participate in the study is a weakness. It is however considered to be sufficient, as the experts are assumed to be more coherent in performance than the novices. For some of the metrics not showing statistically difference, the experts still performed different from the majority of the novice group (Fig. 1). Inclusion of more experts might have resulted in more metrics with statistically difference. Inclusion of surgeons of an intermediate experience level would also have strengthened our study. A higher number of experts and

intermediates should be subjects in future studies before developing specific tests of psychomotor skills to be performed on specific procedures.

In summary, our findings indicate that a good performance during a cholecystectomy includes a low *time* to complete the task, high *bimanual dexterity*, good *motion smoothness* and a relatively higher usage of the grasper compared to the US hook. To quantify the latter finding we defined a new metric, *percentage path length grasper*, which is calculated as the path length of the grasper divided by the path length of both instruments. The experts scored significantly higher than the novices for this new metric.

Even though we were able to find motion analysis metrics that differentiate the performance of experts and novices, and thus did establish construct validity, care should be taken in relying on motion analysis only to assess performance. Motion data will not necessarily reflect errors, e.g. perfusion of the gall bladder and perforations in the liver tissue. If motion analyses are to be used in the performance evaluation of a similar clinical setting, there is still a need for an evaluation by an expert. However, combined with the motion analysis, the effort put into the evaluation could be considerably reduced to e.g. an examination of the tissue for perforations after completion. In this experiment the participants performed a cholecystectomy on a porcine liver model. The results can be transferred to other procedures on different organs. However, the selection of assessment metrics and performance criteria has to be adapted to each procedure individually.

In conclusion, an expert performance during a laparoscopic cholecystectomy measured by instrument motions seems to include a low *time* to complete the task, high *bimanual dexterity*, low measurement of *motion smoothness* of the US hook and a relatively higher usage of the grasper compared to the US hook.

Acknowledgment

Disclosure of interest

Erlend Fagertun Hofstad, Cecilie Våpenstad, Lars Eirik Bø, Thomas Langø, Esther Kuhry and Ronald Mårvik have no conflicts of interests or financial ties to disclose.

References

1. Feldman LS, Sherman V, Fried GM: Using simulators to assess laparoscopic competence: ready for widespread use? Surgery 2004, 135(1):28-42.

2. Subramonian K, DeSylva S, Bishai P, Thompson P, Muir G: Acquiring surgical skills: a comparative study of open versus laparoscopic surgery. Eur Urol 2004, 45(3):346-351; author reply 351.

3. Gallagher AG, Satava RM, Shorten GD: Measuring surgical skill: a rapidly evolving scientific methodology. Surg Endosc 2013, 27(5):1451-1455.

4. Moorthy K, Munz Y, Dosis A, Bello F, Darzi A: Motion analysis in the training and assessment of minimally invasive surgery. Minim Invasive Ther Allied Technol 2003, 12(3):137-142.

5. Aggarwal R, Hance J, Darzi A: Surgical education and training in the new millennium. Surgical endoscopy 2004, 18(10):1409-1410.

6. Mason JD, Ansell J, Warren N, Torkington J: Is motion analysis a valid tool for assessing laparoscopic skill? Surgical endoscopy 2013, 27(5):1468-1477.

7. Hofstad EF, Våpenstad C, Chmarra MK, Langø T, Kuhry E, Mårvik R: A study of psychomotor skills in minimally invasive surgery: what differentiates expert and nonexpert performance. Surgical endoscopy 2013, 27(3):854-863.

8. Chmarra, K. M, Klein, Stefan, Winter d, F. JC, Jansen, Frank-Willem, Dankelman et al: Objective classification of residents based on their psychomotor laparoscopic skills. Surgical endoscopy 2010, 24(5):1031-1039.

9. Oropesa I, Sanchez-Gonzalez P, Lamata P, Chmarra MK, Pagador JB, Sanchez-Margallo JA, Sanchez-Margallo FM, Gomez EJ: Methods and tools for objective assessment of psychomotor skills in laparoscopic surgery. J Surg Res 2011, 171(1):e81-95.

10. Gallagher AG, Ritter EM, Satava RM: Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. Surg Endosc 2003, 17(10):1525-1529.

11.   Smith SGT, Torkington J, Brown TJ, Taffinder NJ, Darzi A: Motion analysis. Surgical Endoscopy And Other Interventional Techniques 2002, 16(4):640-645.

12.   Aggarwal R, Grantcharov T, Moorthy K, Milland T, Papasavas P, Dosis A, Bello F, Darzi A: An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room. Ann Surg 2007, 245(6):992-999.