Cureus
Part of SPRINGER NATURE

Open Access Original Article

Analyzing Procedural Time and Its Relationship to Performance in Microsurgical Training: A Pilot Study

Pablo J. Villanueva [1], Taku Sugiyama [2], Yelena Akelina [3], Hector I. Rodriguez [4], Francisco Salguero [5]

1. Microsurgical laboratory, University of Buenos Aires, Buenos Aires, ARG 2. Neurosurgery, Hokkaido Univerisity, Hokkaido, JPN 3. Simulation, Columbia University College of Physicians and Surgeons, New York, USA 4. Engineering, Salta National University, Salta, ARG 5. Informatics, Universidad Nacional de Cuyo, Mendoza, ARG

Corresponding author: Pablo J. Villanueva, pvillanuevach@gmail.com

Abstract

Introduction

Objective measures and precise variable selection are cornerstones of skill assessment. Currently, the quantity and quality of movements in microsurgery are evaluated using advanced technology and innovative methods. Even so, it is not easy to move the focus out of the idea that improving skills generates faster procedures. Critical analysis reveals that quality does not always align with speed, nor does better performance necessarily mean faster performance. This study aims to analyze and describe the progression of time variables during skill training.

Methods

A published microsurgical training protocol guided this trial, measuring time and errors at each stage. Procedures were recorded chronologically, tracking the number of procedures performed per session and the intervals between sessions. Data were captured using Misat-APP®, exported to a PostgreSQL® database, and processed via Python® scripts to ensure data integrity and facilitate statistical analysis. Results were presented through detailed graphical descriptions.

Results

This trial accounted for 150 protocol procedures, performed over nine months and 75 sessions. The total time spent in this essay was 61h 22min 55sec. The time to complete the procedure presented significant and meaningful fluctuations (slower one = 3082 sec at 4th procedure; faster one = 890 sec, 81st procedure) tracing a decreasing pattern with a significant dispersion level (R^2 0,566). While observing intersession time gaps, strong correlation was evidenced between gaps longer than 10 days and significantly increasing the time to complete the task at the next session (paired T-test, p=0.009). In the final plateau-like phase of the trial, correlation emerged between faster procedures and higher error rates, evidenced by using the concept of “Safe Pace” (Average-0 index ±0.2SD). Other findings suggested that time is also linked to the number of procedures in the same session, the procedure strategy (magnification level, instruments, sequence of steps), final end-product quality, and operator´s attitude.

Conclusion

Progression measurements of time-related variables were analyzed and presented. Key findings included high dispersion levels as a stand-alone skill assessment tool and notable sensitivity to external factors. While assessing experienced operators, over speeding procedures seem to be prone to an increased risk of mistake.

Categories: Medical Education, Medical Simulation, Neurosurgery
Keywords: computer platform for surgical evaluation, surgical skill assessment, microsurgical training, microsurgery simulator, super microsurgery

Introduction

Objective measurements and the selection of key variables are crucial in skill assessment [1,2]. Time taken and performance speed are essential landmarks for evaluating an operator's proficiency in any field [3].

While faster procedures might seem advantageous, they are not necessarily better. Performance must balance speed with the rigor of achieving accurate and high-quality results. On one hand, the time taken for an operator's movements to complete a task, serve as a quantitative measure. On the other hand, criteria for task completion becomes the definition for qualitative assessment.

Movement´s quantity can be assessed by the time to complete a task (TCT), reflecting not only the operator's hand speed but also motion economy, naturally linked with skill level and expertise. TCT values are sensitive but influenced by many factors, thus lacking specificity. The quality can be graded by the final product's appearance or the number of mistakes made during the task, which should be objectively scored by experts or other validated methodology. Multiple studies have investigated this phenomenon using varied tools, techniques, and technologies, each with distinct cost, complexity, and accuracy profiles [4-8].

The speed of an operator is a valuable feature, but this should not overshadow other crucial aspects such as precision, consistency, and overall effectiveness [9]. Achieving a balance ensures comprehensive performance improvement without compromising on essential qualities.

A deep understanding of time-related variables seems to offer a proper background to interpret the complex scenario of hand skill assessment. This pilot trial aimed to register, analyze, and present descriptive details about the time variable during the performance of microsurgical training tasks.

Materials And Methods

Training Protocol

This pilot trial employed a validated microsurgical training protocol designed to objectively evaluate the acquisition of hand skills. This protocol was chosen because it offers clearly defined, progressive microsurgical tasks using human placenta as a realistic simulator for end-to-end vessel anastomosis. A detailed description of the protocol and its associated learning curve analysis can be found in previous publications [10, 11]. Additionally, to minimize bias from biological variability, we incorporated a difficulty grading scale specifically for placenta-simulator quality standardization [12].

Time variables and procedural mistakes

Evaluated records included the time needed to complete the task (TCT) and the time elapsed between training sessions, the inter-session time (IST). When more than one procedure was performed within the same session, there was no significant time between them. Thus, the inter-procedure time will be considered zero (immediate beginning of the next procedure after finishing the previous one).

To objectively assess procedural mistakes, we applied a modified error-scoring system previously reported:

Minor mistake (1 point): A technical error that did not compromise the final outcome and was potentially repairable (e.g., minor leakage during patency testing).

Major mistake (5 points): A critical error that prevented successful completion of the procedure or required substantial corrective action (e.g., arterial wall tear during dissection).

To improve precision while interpreting TCT values, the original exclusion criterion (regarding mistake level at each procedure) was modified from a value of 5pts to 3pts (procedures with a mistake score higher than 3pts were discarded).

However, major mistakes (accounting 5 pts) were still recorded for further interaction with a new index. This index was named “Average-O” (AVO).

$$AVO = \left[\frac{(\#\#TCT) \text{ of proceudres with ZERO mistakes}}{\text{number of procedures with ZERO mistakes}}\#\right]$$

It was simply calculated by getting the mean TCT value of all procedures accounting zero major mistakes. Then, the AVO was used to define three ranges: a central range called Safe Pace (within ± 0.2 SD from AVO benchmark; this parameter was suggested after data analysis (to obtain a range, as wide as possible, but always mistake-free); an upper range called Faster Pace (all values lower than [AVO - 0.2 SD]); and a lower range called Slower Pace (all values higher than [AVO + 0.2 SD]). The Safe Pace range was named after the finding that all procedures withing this range were free of mistakes.

Data Processing

All sample records were registered by using MisatAPP® (powered by CDLAB, Argentina), a dedicated web application which allowed operators to upload required data from their procedures, access their own results live-cast, and receive performance-related reports (Screenshots at Figure 1). Once data was generated and uploaded, it was automatically anonymized, validated, and exported to feed into a formal database PostgreSQL® based. This allowed flexible and reliable interaction (by using Python® codes) between variables, dedicated statistical software, and prepare the scenario for artificial intelligence engaging. The main structure of raw data can be evaluated at Table 1, and a graphic description of data processing at Figure 2.

PROCEDURE	DATE	MISTAKES				TIME				
		360	adventitia	suture	TOTAL	360	adventitia	suture	TOTAL	GAP
1	24-04-18			1	1	776	203	1159	2138	0
2	24-04-19		1	1	2	1002	859	752	2613	1
3	24-04-19				0	408	562	901	1871	0
4	24-04-20		1		0	484	648	1925	3057	1
5	24-04-24				0	619	452	1865	2936	4
6	24-04-25			1	1	690	646	1458	2794	1
7	24-04-25			1	1	404	397	840	1641	0
8	24-04-26			2	1	435	135	1028	1598	1
9	24-04-26				0	446	946	690	2082	0
10	24-04-26				0	519	593	822	1934	0
11	24-04-29				1	246	505	1188	1939	3
12	24-04-30				0	472	508	962	1942	1
13	24-05-02		1	1	2	711	648	1079	2438	2

TABLE 1: Raw data records (sample of 13 procedures).

Sample of 13 procedures data frame, generated by Misat-APP. Variables for mistakes and times were individualized by stages (stage 1 “360”, stage 2 “adventitia”, stage 3 “suture”). A new variable was calculated (from the procedures date values, to measure (in days) the time gap between sessions). The complete raw table can be found in the appendix section. Source: own elaboration.

(A) (B) (C) (D)

FIGURE 1: Cellphone based application for microsurgical data recording. MiSAT-App ® screenshots and Microsoft PowerBI ® dashboard.

(A) Main screen to select working Protocol and User details. (B) Results of a just-ended session (this information, once the procedure ended, is automatically and immediately recorded at the main dataset, MySQL® and Postgres ® compatible). (C) Screenshot of protocol selection, instructions and options. (D) Automatically and live generated Dashboard for all the operators during a session. The number of operators, sessions and dates can be selected to evaluate different stages of a trainee cohort or similar segregation methods. Source: own elaboration.

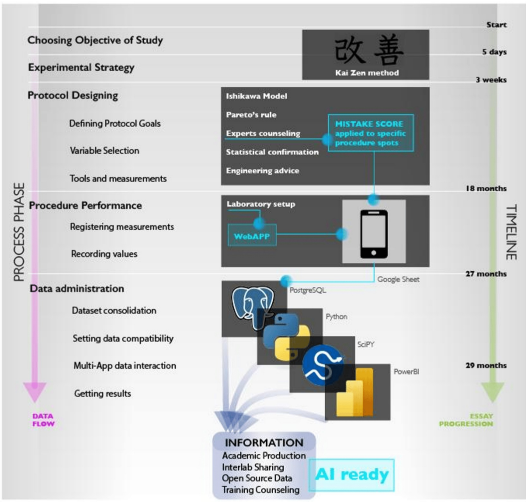

FIGURE 2: Data flow and data processing diagram

This infographic illustrates the data flow from the initial requirement (microsurgical training) to the final goal (multi-purpose information), depicted by phases and time consumed at each phase. These phases were scientifically based, and evidence was generated (training protocols, index and values of interpretation) when not available. The entire system has been verified to be reliable across various scenarios and also being ready for Data Intelligence algorithms. Source: own elaboration.

The complete trial was performed by just one operator, who was previously instructed in microsurgical skills and the performance of this protocol.

Statistical analysis

For overall data: simple counting and logarithmic trend line. Data dispersion was calculated as Coefficient of determination (R2).

For multiple sessions and inter-session analysis: when multiple procedures were conducted within a single session, TCT of first and second procedure were compared using paired T-test (when more than two procedures were found at the same session, no significant difference were found for following procedures after the second one). When the IST was 10 days or longer, the difference between TCT before and after this interval was analyzed. Outliers were identified (±2SD) and removed to improve precision. The correlation between this difference and IST value was analyzed by using paired T-test.

For AVO index: simple average (mean value) and Standard Deviation calculations.

Statistical significance was set at a p-value <.05 for all evaluated procedures.

Results

Overall data and learning curve

This trial accounted for 150 procedures (from a total of 162, twelve of them excluded by meeting the updated Major Mistake criterion). Those were performed for nine months, in 76 sessions, with an average of 1.97 procedures per session (Figure 3).

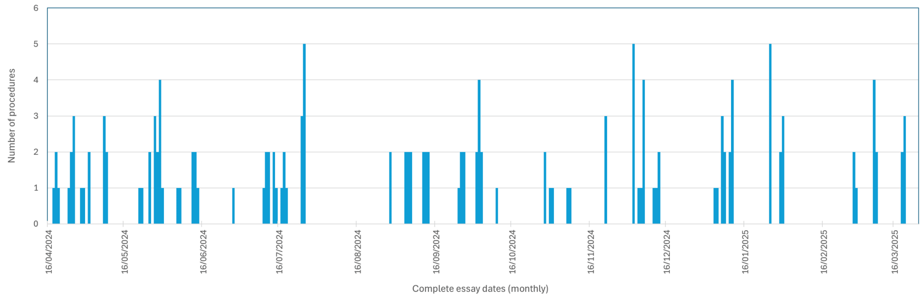

FIGURE 3: Training sessions and their intensity (number of procedures) during the complete trial.

Bars = sessions. Bars height = number of procedures in that session. The horizontal axis shows a monthly chronological progression. Blank spaces between bars represent time-gaps between sessions (days without any microsurgical activity). Source: own elaboration.

The total microscope time for this essay was 61h 22min 55sec. The TCT presented a wide range of values (slower one = 3057 sec at 4th procedure; faster one = 890 sec, 81st procedure), tracing a decreasing trend line (Figure 4) with a significant dispersion level (R2 0.6001).

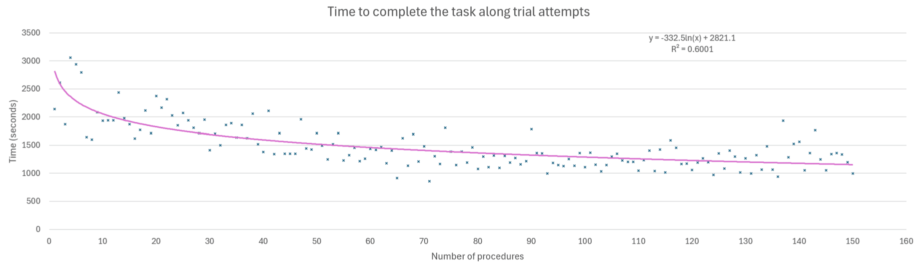

FIGURE 4: Values evolution for the Time to Complete the Task (TCT) variable, with its trend line and scatter dots for each procedure.

The Time to Complete the Task (TCT) shows a downward trend. Blue dots indicate each record in this trial, revealing high dispersion levels (R2 value of 0,601). Source: own elaboration.

Multiple procedures within a single session

Data was segregated by “sessions accounting for more than one procedure”, and 47 sessions accomplished this criterion. After segregating, procedure´s TCT values within the same session were compared. It resulted that 44 from 47 times (93,62%) the second procedure was significantly faster than the first one (Figure 5). The average time of the first and second procedure were 1690.23 seconds, and 1313.83 seconds respectively (Paired T-test, p=0.00001).

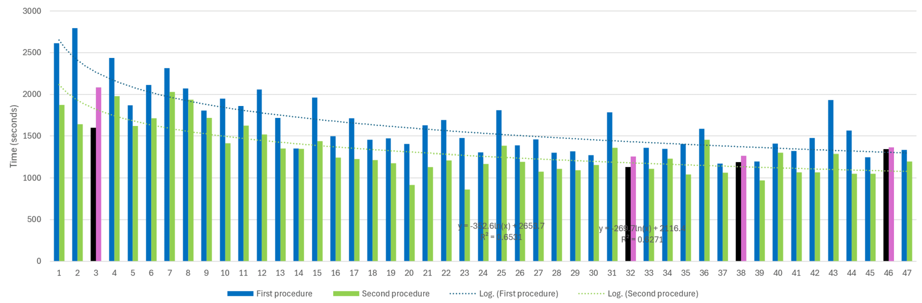

FIGURE 5: Comparison of Time to Complete the Task values (TCT) and respective trend lines, between the first procedure (blue colored) and the second procedure (orange colored), within the same session.

Forty-seven sessions (X axis) were detected to have at least two procedures. TCT presented as vertical bars (blue for first procedure; green for second procedure) showing the most common pattern (blue bars higher than green ones); when this pattern inverted, colors were switched to black and magenta to easier visualization. Trend lines from both remained significantly apart, even at plateau levels. Source: own elaboration.

Inter-session time and microsurgical performance

While observing procedures performance after a significant inter session time (IST), a correlation was found: the first procedure after a long IST had significantly higher TCT levels. By analyzing TCT variations along with the IST levels, the mentioned correlation was evidenced (Figure 6).

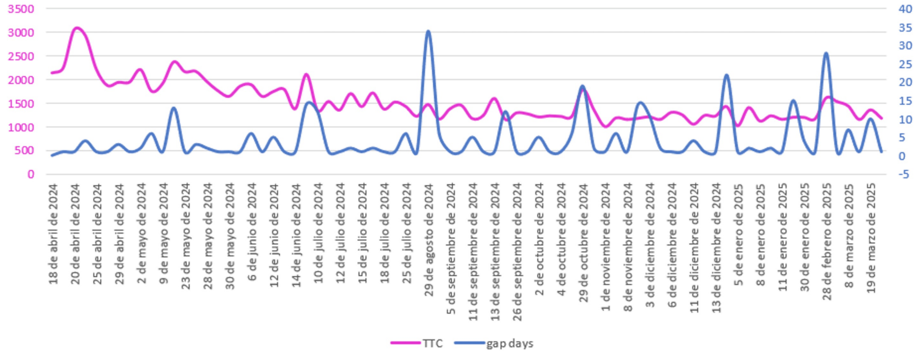

FIGURE 6: Time to Complete the Task (TCT) variations throughout the trial, compared to the Inter Session Time (IST) levels.

Pink line: the TCT values over the complete trial (measured in seconds). Blue line: the fluctuation of IST amounts (measured in days). The peaks of both lines show a tendency to correlate. Source: own elaboration.

During this trial, an IST larger than 10 days was found 8 times. The pair of sessions separated by this large IST was taken, and TCT levels were compared, taking the first procedure of the earlier session and the first procedure of the later session (Figure 7). After this segregation, the results were refined by measuring Standard Deviation of the sample and removing outliers (outlier defined as ±2SD; one was found and removed before statistical analysis).

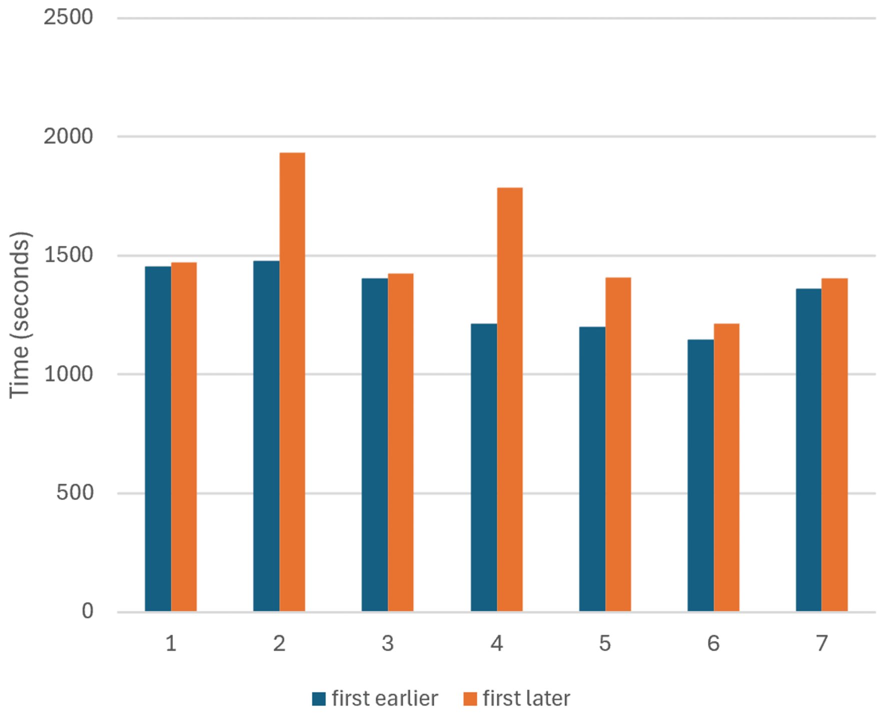

FIGURE 7: Values of Time to Complete the Task (TCT), comparing results of procedures performed before and after an inter session time (IST) longer than 10 days.

Seven pairs of sessions were found to be separated for 10 days or more (after outliers identification (above 2SD range) and being excluded). Blue bars: first procedure of the earlier session. Orange bar: first procedure of the later session. (Paired T-test, p = 0.009). Source: own elaboration.

AVO index

Overall, a descending TCT trend correlated with fewer mistakes, both indicating skill acquisition. However, at the learning curve's plateau, faster procedures resulted in a slight increase in major mistakes. This was revealed by analyzing the complete dataset (prior to applying exclusion criteria, thus including all procedures with a major mistake), which enabled the creation and use of the AVO index (see methodology section) and the construction of Table *2*:

Pace Category	Number of Procedures	Number of Major Mistakes	Percentage of Major Mistakes
Faster (< 1389.94s)	52	1	1.92%
Regular (1389.94s - 1536.61s)	25	0	0.00%
Slower (> 1536.61s)	73	7	9.59%

TABLE 2: Segregation of procedures using the Average-0 index (average time for all procedures with zero major mistakes).

The Average-0 index was used to set three different categories. The central category was created by using 0.2 SD, which also matches 0,5% from benchmark. Thus, Regular Pace was defined as TCT values between 1389.94 and 1536.61 seconds (at this pace the procedure was mistake-free). Source: own elaboration.

As a brief comment to this last table, it could be seen that stepping out of Regular Pace carried a higher risk. Slower Pace accounted for 9.59% of major mistake risk, while Faster Pace accounted for 1.92% of major mistake risk.

Discussion

Continuous improvement-oriented analysis of a procedure combines surgical, technical, and engineering fields [13]. Identifying critical elements within a procedure is crucial for its favourable evolution. These elements then become "improvement opportunities".

Once identified, prioritization based on impact is necessary to enable appropriate monitoring of significant variables. This approach allows a procedure to enter an improvement cycle, where regular, small adjustments lead to substantial positive changes. Numerous authors underscore the significance of knowing precisely where to watch and what to monitor [14, 15].

Overall, for a manual procedure, monitoring necessitates a deep understanding, study, and observation of the quantity and quality of hand movements [16 - 18]. While the quantity of movement can be equated with Time to Complete Task (TCT), this measure can be crude and inaccurate [16], often varying widely due to minor, frequently overlooked differences. This variability is particularly notable in microsurgery, highlighting it as a critical area for analyzing time-related data, including TCT, time between procedures, order of performance, and gap-time of inactivity between sessions, among others.

Overall data

Figure *3* summarizes the trial framework. The X-axis represents chronological order, with regular monthly separations. Bars denote sessions, with their height indicating the number of procedures per session. Blank gaps between bars signify periods of inactivity for the operator.

Time to Complete Task (TCT) correlated well with skill gaining (lower TCT indicating faster procedures). However, additional steps were required to verify and validate this possible fact (e.g., correlating with mistake values, final product quality observations, etc). This suggests that TCT is a sensitive variable but also context dependent. Figure *4* depicts a decreasing trend line (skill correlation) with high dispersion (R2 = 0.601), indicating that TCT should not be used as the sole marker of skill evolution.

Multiple sessions and inter-session

Deeper data analysis revealed 47 sessions with more than one procedure. In 93.62% of these (44 sessions), the first procedure was significantly slower than the second. Figure *5* depicts these 47 sessions, showing separate lines for first and second procedure TCT values. This suggests the initial procedure served as an effective "warming-up" process, potentially leading to more fluid and precise hand movements, among other possible explanations for this effect (e.g., increased relaxation, improved confidence, or better understanding of simulator tissue response).

It was also valuable to incorporate the Inter Session Time (IST), measured in days without microsurgical activity, as a parameter for interpreting microsurgical procedures. Figure *6* illustrates IST, showing higher peaks corresponding to a greater number of inactive days. When TCT levels are plotted chronologically and synchronized with the IST line, both display simultaneous and coordinated peaks. This coordination between IST and TCT peaks clearly demonstrates that days without practice negatively impact procedure performance, resulting in increased task completion times.

The aforementioned situation becomes more solid and statistically confirmed (Paired T-test, p=0.009) with an IST greater than 10 days (Figure *7*).

It was also observed that upon returning to the laboratory after a significant IST, the operator's skill appeared to revert to a nearly-normal level after just a few minutes of training (typically after one completed procedure, approximately 18-25 minutes). This evidence suggests a recommendation for microsurgeons to avoid prolonged periods without training and maintaining at least a regular and simple practice routine.

Average zero index

A different approach involved studying how different procedural speeds affected outcomes. To gain a more accurate understanding of these results, this information was cross-referenced with Major Mistakes (MM), thereby combining quantitative (TCT) and qualitative (MM) evaluation methods.

To investigate the possibility of a particular "pace" or rhythm proving more effective or secure, the Average-0 index was developed. This index identifies an optimal speed for procedures considered most effective and safe (i.e., those with zero major mistakes). Table *2* presents the thresholds defining a Regular zone (safe zone, ± 0.2 SD around Average-0), a Faster zone (below regular zone values, low mistake risk), and a Slower zone (above regular zone, high mistake risk). Several interpretations can be drawn from this finding: a) If a safe pace exists, it should be calculated individually for each operator and each procedure. b) The high risk associated with slower procedures could indicate struggling with an error-prone situation. c) High risk for slower procedures might also suggest a lack of skill or a substantial IST without intervening training. d) The low risk observed for faster procedures could correlate with factors such as overconfidence, insufficient precaution, or an undue emphasis on speed.

Regardless of the underlying reasons, the safe pace clearly provides evidence to recommend that an operator (including young trainees while training) identify and adhere to this rhythm. This approach encourages performing procedures deliberately, without unwarranted rushing or excessive speed, in pursuit of a safe and high-quality outcome.

The findings from Table *2* require further data and corroboration, but they introduce an additional area of inquiry: how time-related variables (time, speed, and pace) could serve as targets for microsurgical skill assessments and procedure evaluations, as well as how to correctly interpret their values.

To complete our paper discussion, the operator's subjective comments should be mentioned. The operator suggested that varying magnification levels and altering instruments or strategies can also influence the Time to Complete Task (TCT), supporting the idea that TCT can be affected by many external factors.

The observations throughout this trial allowed for a thorough description of the progression of time-related variables, providing various insights and nuances relevant to skill training.

Limitations

Additional statistical tests and a larger series of multiple operators (aiming to enhance generalizability, varying operator experience, and to include feedback surveys) are needed to confirm and bring evidence to this essay findings.

Conclusions

Time-related variables were found to be sensitive but lacking specificity as an independent skill marker. This last assertion is supported by TCT high dispersion levels (R2 = 0.6001), and noticeable interaction with external variables (e.g.: IST significant influence; the order within a session; and other annotations from the operator).

Interestingly, the TCT (while being cross-referred to other qualitative variables) presented the ability to identify different paces of performance, showing correlation with quality/safety levels.

The observations throughout this trial allowed for a thorough description of the progression of time-related variables, providing various insights and nuances relevant to skill training.

Appendices

PROCEDURE	DATE	MISTAKES				TIME				
		360	adventitia	suture	TOTAL	360	adventitia	suture	TOTAL	GAP
1	24-04-18			1	1	776	203	1159	2138	0
2	24-04-19		1	1	2	1002	859	752	2613	1
3	24-04-19				0	408	562	901	1871	0
4	24-04-20		1		0	484	648	1925	3057	1
5	24-04-24				0	619	452	1865	2936	4
6	24-04-25			1	1	690	646	1458	2794	1
7	24-04-25			1	1	404	397	840	1641	0
8	24-04-26			2	1	435	135	1028	1598	1
9	24-04-26				0	446	946	690	2082	0
10	24-04-26				0	519	593	822	1934	0

11	24-04-29				1	246	505	1188	1939	3
12	24-04-30				0	472	508	962	1942	1
13	24-05-02		1	1	2	711	648	1079	2438	2
14	24-05-02		1		1	366	573	1039	1978	0
15	24-05-08		1		1	444	403	1020	1867	6
16	24-05-08			1	1	292	338	989	1619	0
17	24-05-08				0	648	473	646	1767	0
18	24-05-09			1	1	484	308	1321	2113	1
19	24-05-09				0	483	273	957	1713	0
20	24-05-22			2	2	477	423	1473	2373	13
21	24-05-23			2	2	384	460	1320	2164	1
22	24-05-26				0	288	548	1478	2314	3
23	24-05-26				0	426	746	856	2028	0
24	24-05-28			1	1	712	194	947	1853	2
25	24-05-28			1	1	791	131	1150	2072	0
26	24-05-28				0	480	240	1216	1936	0
27	24-05-29			1	1	514	302	988	1804	1
28	24-05-29			1	1	481	458	776	1715	0
29	24-05-30			1	1	253	496	1200	1949	1
30	24-05-30		1	1	2	422	340	651	1413	0
31	24-05-30			1	1	260	387	1055	1702	0
32	24-05-30			1	1	203	170	1127	1500	0
33	24-05-31			1	1	745	242	869	1856	1
34	24-06-06			1	1	432	619	838	1889	6
35	24-06-07				0	298	256	1086	1640	1
36	24-06-12			1	1	458	365	1037	1860	5
37	24-06-12			1	1	303	379	942	1624	0
38	24-06-13			1	1	489	620	948	2057	1
39	24-06-13				0	321	240	956	1517	0
40	24-06-14			1	1	299	282	797	1378	1
41	24-06-28			1	2	370	405	1332	2107	14
42	24-07-10			1	2	301	115	927	1343	12
43	24-07-11			1	1	422	284	1009	1715	1
44	24-07-11				0	196	183	971	1350	0
45	24-07-12				0	380	194	777	1351	1
46	24-07-12				0	301	277	769	1347	0
47	24-07-14				0	356	506	1097	1959	2
48	24-07-14				0	285	279	876	1440	0
49	24-07-15				0	404	376	643	1423	1
50	24-07-17			1	1	253	213	1249	1715	2

51	24-07-18			1	1	328	375	793	1496	1
52	24-07-18				0	281	167	798	1246	0
53	24-07-19				0	343	416	761	1520	1
54	24-07-25			1	1	481	150	1082	1713	6
55	24-07-25				1	366	256	606	1228	0
56	24-07-25				0	278	223	825	1326	0
57	24-07-26			1	1	266	253	935	1454	1
58	24-07-26				0	231	298	688	1217	0
59	24-07-26			1	1	220	411	630	1261	0
60	24-07-26			1	1	226	277	932	1435	0
61	24-07-26				0	189	549	677	1415	0
62	24-08-29				0	530	264	676	1470	34
63	24-08-29				0	224	370	582	1176	0
64	24-09-04			2	2	326	385	692	1403	6
65	24-09-04				0	204	138	573	915	0
66	24-09-05				0	302	442	882	1626	1
67	24-09-05			1	1	270	225	637	1132	0
68	24-09-06			1	1	306	376	1010	1692	1
69	24-09-06				0	200	238	773	1211	0
70	24-09-11				0	298	440	740	1478	5
71	24-09-11				1	203	159	498	860	0
72	24-09-12			1	1	310	378	614	1302	1
73	24-09-12				0	243	331	594	1168	0
74	24-09-13				0	243	572	995	1810	1
75	24-09-13				0	301	238	845	1384	0
76	24-09-25				0	131	192	824	1147	12
77	24-09-26			1	1	225	278	885	1388	1
78	24-09-26				0	185	346	661	1192	0
79	24-09-27			1	0	249	194	1016	1459	1
80	24-09-27				0	123	179	776	1078	0
81	24-10-02			2	2	384	232	682	1298	5
82	24-10-02			1	1	206	210	693	1109	0
83	24-10-03			1	1	266	307	743	1316	1
84	24-10-03			1	1	330	189	575	1094	0
85	24-10-03				0	445	140	723	1308	0
86	24-10-03				0	335	87	769	1191	0
87	24-10-04				0	214	147	911	1272	1
88	24-10-04			1	1	170	281	706	1157	0
89	24-10-10			1	1	196	192	827	1215	6

90	24-10-29			3	3	270	412	1103	1785	19
91	24-10-29			1	1	353	193	814	1360	0
92	24-10-31				0	243	319	792	1354	2
93	24-11-01				0	179	102	718	999	1
94	24-11-07				0	231	230	723	1184	6
95	24-11-08			1	1	182	230	737	1149	1
96	24-11-22			1	1	237	328	565	1130	14
97	24-11-22				0	200	297	760	1257	0
98	24-11-22				0	280	85	771	1136	0
99	24-12-03			2	2	279	249	832	1360	11
100	24-12-03			1	2	269	57	784	1110	0
101	24-12-03			1	2	156	224	988	1368	0
102	24-12-03			1	2	184	111	858	1153	0
103	24-12-03				0	213	145	678	1036	0
104	24-12-05				0	156	211	783	1150	2
105	24-12-06			1	1	261	235	805	1301	1
106	24-12-07			1	1	184	284	878	1346	1
107	24-12-07				0	230	205	797	1232	0
108	24-12-07				0	183	292	730	1205	0
109	24-12-07			1	1	283	192	728	1203	0
110	24-12-11				0	222	302	525	1049	4
111	24-12-12				0	298	276	661	1235	1
112	24-12-13			1	1	273	208	922	1403	1
113	24-12-13				0	250	196	597	1043	0
114	25-01-04			1	1	424	282	718	1424	22
115	25-01-05				0	186	179	651	1016	1
116	25-01-07				0	243	241	1101	1585	2
117	25-01-07			1	1	230	315	909	1454	0
118	25-01-07			1	1	291	206	672	1169	0
119	25-01-08			1	1	217	282	674	1173	1
120	25-01-08				0	252	157	653	1062	0
121	25-01-10				0	357	234	601	1192	2
122	25-01-10			1	1	126	352	787	1265	0
123	25-01-11	1		1	2	481	73	645	1199	1
124	25-01-11				0	215	133	622	970	0
125	25-01-11				0	571	138	646	1355	0
126	25-01-11				0	229	61	795	1085	0
127	25-01-26				0	249	306	853	1408	15
128	25-01-26				0	240	318	740	1298	0
129	25-01-26				0	228	203	587	1018	0

130	25-01-26			2	2	178	321	765	1264	0
131	25-01-26			1	1	178	247	574	999	0
132	25-01-30				0	350	252	720	1322	4
133	25-01-30				0	237	158	671	1066	0
134	25-01-31				0	322	268	888	1478	1
135	25-01-31				0	187	217	664	1068	0
136	25-01-31			1	1	228	96	619	943	0
137	25-02-28			1	1	345	377	1210	1932	28
138	25-02-28			1	1	347	269	670	1286	0
139	25-03-01			1	1	422	263	842	1527	1
140	25-03-08				2	412	213	940	1565	7
141	25-03-08				2	226	185	640	1051	0
142	25-03-08				1	182	314	862	1358	0
143	25-03-08				1	350	190	1224	1764	0
144	25-03-09				1	239	151	858	1248	1
145	25-03-09				0	285	140	625	1050	0
146	25-03-19			1	1	263	181	899	1343	10
147	25-03-19			1	1	213	290	861	1364	0
148	25-03-20			1	1	273	223	837	1333	1
149	25-03-20			1	1	215	244	738	1197	0
150	25-03-20				0	168	187	643	998	0

TABLE 3: Complete Dataset

The complete trial records are presented in this Table. Source: own elaboration.

Additional Information

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Andersen SAW, Nayahangan LJ, Park YS, Konge L: Use of Generalizability Theory for Exploring Reliability of and Sources of Variance in Assessment of Technical Skills: A Systematic Review and Meta-Analysis. Acad Med. 20211, 96:1609-1619. 10.1097/ACM.0000000000004150
2. Asif H, McInnis C, Dang F, et al.: Objective Structured Assessment of technical skill (OSATS) in the Surgical Skills and Technology Elective Program (SSTEP): Comparison of peer and expert raters. The. American Journal of Surgery. Volume. 2021, 223:276-279. 10.1016/j.amjsurg.2021.03.064
3. Niveditha M, Ravi S, Ashish S: Microsurgical suturing assessment scores: a systematic review . Neurosurg Rev. 2022, 45:119-124. 10.1007/s10143-021-01569-3
4. Sugiyama T, Lama S, Gan LS, Maddahi Y, Zareinia K, Sutherland GR: Forces of tool-tissue interaction to assess surgical skill level. JAMA Surg. 2018, 153:234-242. 10.1001/jamasurg.2017.4516
5. Sugiyama T, Sugimori H, Tang M, et al.: Deep learning-based video-analysis of instrument motion in

microvascular anastomosis training. Acta Neurochir (Wien. 2024, 166:6. 10.1007/s00701-024-05896-4
6. Tang M, Sugiyama T, Takahari R, et al.: Assessment of changes in vessel area during needle manipulation in microvascular anastomosis using a deep learning-based semantic segmentation algorithm: A pilot study. Neurosurg Rev. 2024, 47:200. 10.1007/S10143-024-02437-6
7. Lavanchy J, Zindel J, Kirtac K, Twick I, Hosgor E, Candinas D, Beldi G: Automation of surgical skill assessment using a three-stage machine learning algorithm. Sci Rep. 2021, 4:5197-10. 10.1038/s41598-021-88175-x
8. Soangra R, Sivakumar R, Anirudh E, Reddy Y, John E: Evaluation of surgical skill using machine learning with optimal wearable sensor locations. PLoS One. 2022, 3:0267936. 10.1371/journal.pone.0267936
9. Sugiyama T, Ito M, Sugimori H, et al.: Tissue Acceleration as a Novel Metric for Surgical Performance During Carotid Endarterectomy. Oper Neurosurg (Hagerstown. 2023, 1:343-352. 10.1227/ons.0000000000000815
10. Villanueva P, Sugiyama T, Villanueva B, Rodriguez H, Arciénaga A, Cherian I: Using Engineering Methods (Kaizen and Micromovements Science) to Improve and Provide Evidence Regarding Microsurgical Hand Skills. World Neurosurg. 2024, 189:380-390. 10.1016/j.wneu.2024.06.075
11. Villanueva PJ, Rodriguez HI, Sugiyama T, O'Keeffe D, Villanueva G, Villanueva BM, Roche AF: Modeling the Microsurgical Learning Curve Using a Poisson-Based Statistical Approach for Skill Assessment. Cureus. 2025, 25:83009. 10.7759/cureus.83009
12. Villanueva P, Villanueva B, Arcienaga A, Rodriguez H, Lausada N, Lagier M. (2024: A novel scoring system to reduce bias in placental microsurgery training. 395-411. 10.56238/sevened2024.025-032
13. Villanueva P, Villanueva BM, Sanmarco C, et al.: Kaizen: Engineering tools for development, evaluation, certification and continuous improvement of microsurgical abilities and procedures. Seven Editora. 10.56238/medfocoexplconheci-049
14. Imai, Masaaki. Kaizen: The Key to Japan's Competitive Success. New York: McGraw-Hill Education, 1986.
15. Lizarelli FL, Antony J, Suarez M, et al.: Analysis of the impact of Kaizen practices on ESG performance and the mediating role of digital systems. Business Process Management Journal. 2025, 31:148-175. 10.1108/bpmj-01-2025-0056
16. Datta V, Mackay S, Darzi A, Gillies D: Motion analysis in the assessment of surgical skill . Comput Methods Biomech Biomed Engin. 2001, 4:515-523. 10.1080/10255840108908024
17. Ghasemloonia A, Maddahi Y, Zareinia K, Lama S, Dort JC, Sutherland GR: Surgical skill assessment using motion quality and smoothness. Journal of Surgical Education. 2016, 74:295-305. 10.1016/j.jsurg.2016.10.006
18. Azari DP, Frasier LL, Quamme SRP, et al.: Modeling surgical technical skill using expert assessment for automated computer rating. Annals of Surgery. 2017, 269:574-581. 10.1097/sla.0000000000002478