Usability Test for an Expandable Interbody Fusion Cage

Jaemin Yu

Department of Medical Device Industry
The Graduate School, Yonsei University

Usability Test for an Expandable Interbody Fusion Cage

Directed by Professor Sung Uk Kuh and Won Seok Jang

A Master's Thesis
Submitted to the Department of Medical Device Industry
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Master of Engineering

Jaemin Yu

December 2023

This certifies that the master's thesis of
Jaemin Yu is approved.

[signature]

_____

Thesis Supervisor: Sung Uk Kuh

[signature]

_____

Thesis Committee Member: Won Seok Jang

[signature]

_____

Thesis Committee Member: Hyun Seung Yu

The Graduate School
Yonsei University

December 2023

Acknowledgements

With the consistent encouragements and supports from my professors and colleagues, I have gained an opportunity for substantial personal growth during my academic journey.

First and foremost, I would like to express my gratitude to Professor Sung Uk Kuh and Won Seok Jang, who has been consistently helpful from the beginning. With their guidance, I was able to navigate the challenges and successfully complete my academic journey. I also appreciate Professor Hyun Seung Yu for dedicating his time to provide guidance and support in completing my thesis.

Special thanks to the members of Gangnam Severance Hospital Medical Device Usability Center whose collaborative efforts were crucial in the successful completion of the usability test.

Lastly, my heartfelt appreciation goes to Kang U Kim, my companion throughout this academic journey. His shared experiences and mutual support have added depth and richness to my educational endeavors.

Table of Content

List of Tables

Abstract

Usability Test for an Expandable Interbody Fusion Cage

Department of Medical Device Industry
The Graduate School, Yonsei University

(Directed by Professor Sung Uk Kuh and Won Seok Jang)

This paper provides a detailed account of a summative usability evaluation conducted on a newly developed expandable cage designed for lumbar interbody fusion surgery. Lumbar interbody fusion is a widely used surgical intervention for addressing degenerative lumbar spinal diseases. The expandable cage under evaluation represents an advancement in lumbar fusion technology, aiming to minimize iatrogenic damage during insertion. As the newly developed device differs not only from its effectiveness but also from its user interface, it is important to evaluation the usability of the device.

The study employed a comprehensive usability test involving 15 spinal surgeons, who are the intended users of the expandable cage. The test incorporated use-scenarios including critical tasks derived from the risk management profile, and participant feedback to assess various aspects.

Out of a total of 24 tasks, there is one task (Task 17) with a success rate of 73% which was below the criteria, 80%. The rest of the tasks achieved a success rate of 80% or higher, meeting the initial goal. The specific completion results are shown by the table 2. In the survey regarding convenience and safety for each of the 24 tasks, all tasks scored an average of 3.0 points or higher. The convenience score averaged above 4.0 points, while the safety score averaged above 4.2 points. For a total of 11 satisfaction survey items, all items scored an average of 3.0 points or higher. Each satisfaction item received an average score of 4.4 points or higher.

Despite some limitations regarding the test, the usability test provided enough indication of the usability of the device as well as the direction for further improvement. It highlights the significance of usability testing as a crucial step in ensuring the safety and effectiveness of medical devices, offering valuable insights for further enhancements and future iterations of the expandable cage.

In conclusion, while the tested expandable cage demonstrated a proper level of usability and participant satisfaction. Continuous improvement based on user feedback and observed use-errors is crucial for enhancing the overall safety and efficacy.

Key Words: expandable cage, expandable interbody fusion cage, usability, usability test, medical device

Usability Test for an Expandable Interbody Fusion Cage

Department of Medical Device Industry
The Graduate School, Yonsei University

(Directed by Professor Sung Uk Kuh and Won Seok Jang)

I. Introduction

1.Degenerative Lumbar Spinal Disease & Lumbar interbody fusion operation

Degenerative lumbar spinal disease (DLSD) stands out as a significant healthcare challenge on a global scale. This term encompasses a range of conditions, including disc degeneration, lumbar spinal stenosis, and spondylolisthesis, forward displacement of one vertebra with respect to an adjacent vertebra (Ravindra et al. 2018; Katz et al. 2022). The annual global incidence of adult DLSD is estimated at approximately 266 million (Ravindra et al. 2018).

The typical approach for severe DLSD involves combining decompression with lumbar arthrodesis or interbody fusion. The lumbar interbody fusion (LIF) procedure involves fusing adjacent vertebrae to restrict motion, aiming to relieve pain associated with spine instability. Fusion is typically achieved using autologous or allogenic bone grafts in between vertebral bodies, and is often complemented with instrumentation like screws, rods, and/or interbody cages. Interbody cages are devices designed to contain bone graft and inserted into the intervertebral space, the space between adjacent vertebrae, to promote fusion. Surgeons may utilize various approaches for insertion of interbody fusion material, namely anterior, lateral, or posterior (Katz et al. 2022).

With improvements in technology and the design of interbody cages, the rate of lumbar fusions performed has increased over the past decades. However, surgical complications, particularly post-LIF non-union or pseudarthrosis, remain a major challenge. The rates of non-union after LIF still ranged from 7 to 20% with a significantly higher incidence in

cases spanning 3 or more spinal levels. Autogenous iliac crest bone grafting has been used as the gold standard to enhance bony fusion of affected intervertebral joints. However, the harvesting of autologous bone is associated with significant risk of morbidity, intraoperative complications, and donor site pain. Spine surgeons and researchers have been working together to improve fusion rates.(Baliga, Treon, and Craig 2015)

2. Expandable cage

Expandable cages refer to interbody cages that can be inserted with a reduced profile and expanded in situ, promoting fusion and disc height restoration. The use of expandable interbody fusion cages in spinal surgery represents an advancement over static devices, aiming to minimize iatrogenic damage on vertebral body during insertion. Studies indicate that, compared to static cages, expandable cages have advantages in disc height expansion and lordotic restoration, leading to further decompression of nerve root and stabilization of spine. The implants can be inserted into the intervertebral space in a shrunken form and expand, eliminating the need for other maneuvers to expand the disc space. This approach may reduce trauma to the endplate and minimize implant subsidence. Expandable cages offer a promising option in spinal surgery, providing benefits in terms of reduced endplate damage and improved fusion outcomes. (Macki et al. 2021)

Existing expandable cages have various mechanisms of expanding the height of the cage and fix it steady. Thus, different expandable cages have different interactions with users and usability plays important role in developing and utilizing the device (Lewandrowski, Ferrara, and Cheng 2020). This paper will review the summative usability evaluation of a newly developed expandable cage.

3. Human Factor Engineering / Usability Engineering

Medical device use errors, often caused by poorly designed user interfaces, are a major risk for patient harm and death (Shin and Lee 2023). Evaluating the user interface is crucial, as a well-designed interface can minimize complications; conversely, use-related errors may pose harm to users and patients (Karsh 2004). Use-related errors can be caused by medical device that was designed not considering the use-context such as use environment, user limitation. The significance of Human Factor Engineering or Usability Engineering (HFE/UE) in the development of medical devices is increasingly being acknowledged. Many countries mandate usability testing as a requirement for medical device approval (Shin and Lee 2023; FDA February 3, 2016).

Many standard and guidance emphasize the usability or human factors for the medical device, especially IEC standard for medical device usability (IEC/ISO, 2015; IEC/ISO, 2016) suggests the Usability Engineering Process to develop the medical device equipped with usability. IEC Technical Report 62366-2 describes the usability methods applicable to each development stage. (Pelayo, Marcilly, and Bellandi 2021) Although the quality of usability testing is highly dependent on the relevant usability test procedure, there are few guidelines on how to develop usability validation procedures for medical devices(FDA February 3, 2016).

4. International Standards Related to Medical Device Usability

A. IEC 62366-1:2015/Amd 1:2020

This standard demonstrates enhanced alignment with safety-related ISO 14971:2019 and provides guidelines. In addition to safety, it also outlines usability engineering methods, with a focus on the application of usability engineering processes. It aims to efficiently reduce hazards through formative and summative evaluations based on user interface-related usage scenarios. It covers aspects related to safety, as well as relevance to factors such as task accuracy, completeness, efficiency, and user satisfaction. (IEC/ISO 2015)

B. IEC/TR 62366-2:2016

It similarly provides guidelines for specific areas useful in performing usability processes apart from safety. This standard serves the purpose of offering information as guidelines and recommendations, not for regulatory or requirement purposes. It provides information not only on efficient methods for complying with the requirements in IEC 62366-1:2020 but also additional information related to usability, including operational efficiency and user satisfaction(IEC/ISO 2015).

C. FDA Human Factors Guidance

The FDA Human Factors Guidance encourages manufacturers to evaluate potential use errors and hazards related to the use of medical devices in order to minimize them. This guidance does not enforce legal or regulatory requirements unless specific regulatory or legal requirements are cited, but it carries a recommendatory nature. In addition to this, as per international standards related to medical device usability, a usability engineering plan and execution are required. Activities conducted in accordance with the usability engineering process requirements should be documented and submitted in the usability engineering file (FDA February 3, 2016).

5. Objective

The usability test was conducted to demonstrate that intended users could use the expandable intervertebral fusion cage without significant errors, confirming the device's safe usability. Based on the test results, this paper will discuss the significance of usability and the evaluation process. The test involves user interaction with the device user interface (UI) based on use scenarios. Test participants engaged with the device UI, and observations were made by test personnel to identify any use-related issues. Following the interaction, surveys or interviews were conducted to identify the causes of use-related problems or gather data for commercial considerations.

II. Materials and methods

1. The Test Participants

The test involved participants from the intended user group, spinal surgeons. The American National Standard recommends a minimum of 15 participants for a summative evaluation, and accordingly, 15 spinal surgeons were recruited (IEC/ISO 2015; FDA February 3, 2016). None of these participants were involved in the consultation or development process for the device, and they had no personal interests in it. To maintain confidentiality, participant information was coded, from P1 to P15. Before initiating the test, all participants were required to complete a consent form, which included agreement for video recording. Following the conclusion of the test, participants received compensation for their involvement.

2. The Test Environment and Equipment

The medical device subjected to the test was an expandable cage named Excender (CG Bio, Seoul, Korea). Test participants were equipped with the implant, the necessary insertion tools, and a user manual. The test involved simulation of the device insertion procedure, performed on a dummy spine model in a room designed to resemble an operating room. This simulated-use testing took place at Gangnam Severance Hospital Usability Center in Seoul, Korea. The test room replicated the conditions of an actual operating room with comparable brightness, humidity, and temperature.

3. The Test Procedure

The test was conducted with one participant at a time, following the procedures outlined in Table 1. Each participant received explanations of the test procedure prior to the actual test.

Procedure	**Description**
Preparation	Test administrators check if the device normally operates before participant arrive.
Introduction	Test administrators receive participation consent from participants

	Test administrators begin video recording then explain purpose, the planned test session activities.
Present the prompt	Test administrators provide the prompt and explain it.
Perform the use-scenario	While participants use the device in accordance with the use-scenario, test administrator observe and record each use error, close call, request for assistance.
Conduct assessment & interview	After completing the use-scenario, participants proceed to assess the satisfaction with performing each task. Simultaneously, test administrator conduct interview to identify root cause of the use problem observed.
	Once participants complete the satisfaction assessment, test administrator conduct the survey on the safety and training

Table 1. The test procedure

4. Critical Task Identification and Use-scenario

As part of their design controls, manufacturers conduct a risk analysis that comprehends the risks associated with device use. The information from the risk analysis is utilized to categorize tasks based on the severity of potential harm resulting from use errors. This categorization helps identify critical tasks, which, if performed incorrectly or not at all, could cause serious harm to the patient or user. The list of critical tasks is employed to shape the use-scenario in the usability test, ensuring it encompasses tasks relevant to the safety and effectiveness of device use (Russ and Saleem 2018; FDA February 3, 2016; IEC/ISO 2019).

The use-scenario, including the critical tasks, is structured in a logical order to emulate a natural workflow. Tasks that naturally follow a logical sequence during device usage are grouped together. The details of the use-scenario are shown in the table 2. During the test, conductors gather observational data while participants execute tasks from the scenario and categorize their performance into three types:

- Task Completion (C): Successfully completing the task without observed or recorded errors or near-errors.
- Completed with Issues (CI): Near-error situations or difficulties encountered during task execution, even though the task is ultimately performed correctly. Users express any difficulties they faced.
- Not Complete (NC): Participants request assistance, fail to complete the task within 120 seconds, or make user errors leading to unexpected device reactions contrary to the manufacturer's or user's expectation

The completion or success rate was calculated as a proportion of the number of participants who completed the task to the total number of participants. The target completion rate and error rate for each task was set at 80% or higher, and 20% or lower respectively.

5. Task Specific Surveys, Satisfaction Assessment, and Interviews

After completing the test, the participants provided feedback by responding to a 5-point scale survey evaluating the safety and convenience of each task. Additionally, participants responded to a separate satisfaction survey questions in table 3, on a 5-point scale as well. The target average score for each task and question was set at 3.5. Finally, the administrator conducted interviews with participants regarding tasks where difficulties arose and that were not completed. During the interview, the participants freely shared their opinions and comments about the device use.

III. Result

Participants received training through video instructions provided by the manufacturer and had hands-on experience using the test device. They were guided through a series of processes involving the use of the test device, including unpacking the implant, measuring the implant size, and inserting the implant with bone graft safely. Subsequently, participants were requested to evaluate their satisfaction with the usability

of the device through a survey. They also provided personal opinions based on their experiences using the test device and similar devices during interviews with the facilitator.

- Task completion: Out of a total of 24 tasks, there is one task (Task 17) with a success rate of 73% which was below the criteria, 80%. The rest of the tasks achieved a success rate of 80% or higher, meeting the initial goal. The specific completion results are shown by the table 2.
- Task-Specific Surveys: In the survey regarding convenience and safety for each of the 24 tasks, all tasks scored an average of 3.0 points or higher. The convenience score averaged above 4.0 points, while the safety score averaged above 4.2 points.
- Satisfaction assessment: For a total of 11 satisfaction survey items, all items scored an average of 3.0 points or higher. Each satisfaction item received an average score of 4.4 points or higher.

Use-Scenario	**Task**		**Completion**			
	#	**Description**	**C**	**CI**	**N**	**R (%)**
Preparation of the implant	1*	Identify the effective date of the implant	15	0	0	100
	2	Identify the size of the implant	15	0	0	100
	3*	Verify the “disposable” statement on the package	12	0	3	80
	4*	Verify sealing of the implant and “sterility” of it	13	2	0	100
	5*	Open the pouch and aseptically drop the implant on the field	15	0	0	100
Measurement of the implant	6	Connect torque limited handle and lateral quick handle with measurement distractor	12	1	2	87
	7	Rotate the torque limited handle to expand measurement distractor and check the visual gauge 12	15	0	0	100
	8	Turn the torque limited handle reversely and shrink the measurement distractor and remove it	15	0	0	100
Insertion of the implant	9*	Place the implant on the implant holder and connect it	14	1	0	100
	10	Verify the measurement module indicator at gauge “0”	13	0	2	87
	11*	Connect the measurement module on the implant holder until it clicks	12	3	0	100
	12*	Insert the implant driver in the implant holder	15	0	0	100

	13	Combine the torque limited handle and the lateral quick handle with the implant holder	12	1	2	87
	14*	Place the lateral quick handle parallel to the coronal plane, and insert the implant (may use the mallet)	15	0	0	100
	15*	Turn the torque handle to expand the implant until the clicking sound	15	0	0	100
	16	Press the release button to remove the measurement module	15	0	0	100
	17	Insert the bone graft cannula inside the implant holder and use the bone graft pusher to fill in the implant	9	2	4	73
	18	Remove the bone graft cannula from the implant holder	14	1	0	100
Tightening of the locking screw	19*	Connect the torque limited handle with the locking screwdriver	12	1	2	87
	20*	Fit the locking screw on the locking screwdriver	13	2	0	100
	21*	Insert the locking screw drive into the implant holder.	15	0	0	100
	22*	Turn the torque limited handle to fasten the screw until it clicks	10	5	0	100
	23*	Remove the locking screwdriver from the implant holder	15	0	0	100
	24*	Unwind the wheel of the implant holder to remove it from the implant	13	2	0	100

Table. 2 Use-scenario and completion status
The tasks with asterisk(*) indicate critical tasks. # = number, C = Completed, CI = completed with issue, NC = Not completed, R =completion rate

Failure in the tasks

Five task failures were observed in two critical tasks: task 3 and 19. Three users, P1, P3, and P8, were unable to complete the task requiring the identification of the disposability of the device. The risk analysis highlights the potential hazards associated with the reuse of the device, including the possibility of infection and inflammation, particularly spondylitis, which poses a serious threat to patient. Two users, P7 and P8 encountered difficulties in correctly utilizing the locking screwdriver while fastening the locking screw on the cage. In both instances, the users mistakenly attached the torque-limited handle to the implant screwdriver instead of the appropriate locking screwdriver. Failures

in tightening the locking screw can result in reversal of implant expansion and possible dislocation of it.

Other failures occurred in non-critical tasks. In task 6, P9 and P14 exhibited an incorrect coupling direction of the Lateral Quick Handle, lacking awareness that the handle should be coupled laterally. In task 10, P12 and P14 failed to confirm that the scale on the Measurement Module was at 0 initially, leading to an inability to proceed with the task. They realized the issue later. In task 13, P9 and P14 displayed an incorrect coupling direction of the Lateral Quick Handle, lacking awareness that the handle should be coupled laterally. In task 17, which was the only one task which did not meet the target success rate, four participants were not able to finish the task correctly. Participants 1, 2, and 14 had difficulty perceiving that the bone graft cannula was not fully attached, while participants 9 were unaware of the short bone graft pusher (pusher-S) and only used the long bone graft pusher (pusher-L) when pushing the bone graft.

Task Specific Survey, Satisfaction Assessment, and Interviews

Surveys were conducted using a 5-point scale to assess the convenience and safety of each task in the use-scenario, and the results from each participant were aggregated. Scores ranged from 1 (strong disagreement) to 5 (strong agreement). The average convenience score across all 24 task-specific surveys was 4 points, and safety received an average score 4.2 points, surpassing the target score of an average of 3.0 points. A satisfaction survey for participants regarding the test device was also conducted using a 5-point scale. The average satisfaction score across all 11 satisfaction survey items was above 4.4 points, exceeding the target score of an average of 3.0 points. The Specific context of the interviews conducted with the participants are provided in discussion section.

#	Description	min score	max score	avg score

1	The provided user manual is clear and understandable.	4	5	4.5
2	The font style and size in the user manual are considered suitable.	2	5	4.4
3	Labeling of information such as expiration date and specifications on the packaging is appropriately placed.	4	5	4.7
4	The sealing and opening status of the packaging is easy to verify.	3	5	4.5
5	The force required to expand the Measurement Distractor is appropriate.	4	5	4.7
6	Confirmation of the expanded height through the visual gauge of the Measurement Distractor is straightforward.	4	5	4.6
7	Distinguishing between the medial and lateral sides of the implant is easy.	2	5	4.5
8	Checking whether the implant is attached to the holder properly is easy.	3	5	4.7
9	Perceiving the clicking sound when turning the torque limiting handle after the implant reaches maximum expansion is easy.	2	5	4.4
10	Confirmation of the expanded height through the Measurement Module is straightforward.	2	5	4.5
11	Perceiving the clicking sound when turning the torque limiting handle after fully locking the Locking Screw is easy.	2	5	4.5

Table 3. Satisfaction assessment
= number, min = minimum, max = maximum, avg = average

IV. Discussion

HFE/UE involves assessing the device user system, which consists of device users, device use environments, and device user interfaces. The outcome of the interaction among these components can lead to either correct use or use error (FDA February 3, 2016).

Device users are the intended users of the device such as surgeons, physicians, nurses, patients, family members, installer, maintenance staff member, and disposer. User characteristics such as functional capabilities, experience and knowledge levels can impact the safe and effective use of the device. Device use environments are the

environments in which medical devices are used. They can be clinical or non-clinical settings including hospitals, surgical suites, home, emergency use, public use, etc. Device user interface includes all interaction points between the user and the device; from device setup (e.g., unpacking, calibration) to device usage, and maintenance activities (e.g., cleaning, battery replacement, repairs). Device user interface comprises the size and shape of the devices, components and accessories, controls, visual displays, auditory and tactile feedback, alarms and alerts, logic and sequence of operation, labeling, and training etc. (FDA February 3, 2016)

Summative evaluation of usability testing is conducted to demonstrate that the device can be used properly under the expected use conditions. The testing should be comprehensive, capturing use errors caused by the design of the user interface, and the results should be applicable to actual use. The testing design should include participants representing the intended users, performing all critical tasks with the final design of the device user interface under realistic conditions. To optimize use safety and effectiveness, the testing must be sensitive to user interface design problems, whether users are aware of errors or not, and should reveal no use errors that could lead to serious harm and could be addressed through user interface design modifications (FDA February 3, 2016).

The User

In summative usability test, the FDA guideline recommend at least 15 test participants from intended user population, which was “spinal surgeons” for the expandable cage (FDA February 3, 2016; Virzi 1992). Out of the 15 participants, only one was female. All of them were belonged to the neurosurgical department except for one, who was an orthopedic surgeon. The average duration of the participants' surgical experiences was 8.2 years, ranging from 1 year to 24 years. Among the participants, five had prior experience with the testing device in either real or virtual environments, and nine participants had experience with similar devices.

The anticipated users of the device, more specifically, are those who already have experiences with traditional static cages. It is possible for the experienced users to make more user errors when using a newly developed device that are different from the devices they are familiar with. Experienced users may be inclined to pay less attention to instructions regarding cage insertion, relying more on their intuition and existing knowledge. On the other hand, novice clinicians who have no or little experience with interbody cages, would follow the guidelines and instructions more thoroughly and be more carefully in executing the tasks. In fact, P14, with only one year of clinical experience, made 2 minor use errors during the test procedure which was below the average number of errors. However, when the participants already have the experience with the exact device, the performance level was higher and tend to make fewer errors in critical tasks. Likewise, the characteristics of the user can impact the test results. Therefore, obtaining qualitative information from the subjects' use-experience beyond observational data, such as surveys and interviews, is crucial.

The Environment

The settings under which simulated-use testing is conducted should be sufficiently realistic so that the results of the testing are generalizable to actual use. To the extent that environmental factors might affect users' interactions with elements of the device user interface, they should be incorporated into the simulated use environment (e.g., dim lighting, multiple alarm conditions, distractions, and multi-tasking) (FDA February 3, 2016). For the test procedure, dummy spinal model was used, and the participants inserted the implant into the dummy model. Due to the limitations of the dummy model, certain details, such as the forces required to insert the implant in real patients or the effect of blood on the visualization of vertebral spaces, may differ from the actual clinical setting. However, the test environment was realistic enough to simulate the actual steps and procedures of the implant insertion.

The Interface

In the task 3, the critical task with errors, the participants experienced difficulties in recognizing the “disposable” label in the package. The error is associated with visibility in the package. As it is a usability limitation with the package rather than the device itself, it is expected to be correctable with minor improvements. The only task below the target complete rate, task 17, involves attaching the bone graft cannula and filling bone graft. Among participants who failed in this task, a majority encountered difficulties in confirming the attachment status of the bone graft cannula. They either failed to verify whether the cannula was fully attached or faced issues such as the cannula falling off during the filling process, leading to the inability to complete the task. Additionally, one participant filled the bone graft without using the bone graft pusher-S first, and relied only on bone graft pusher-L. They seemed to overlook or did not remember the content in the instructional video that demonstrated the use of bone graft pusher correctly, resulting in the omission of some steps. In task 19, two subjects attached the torque-limited handle to the implant screwdriver instead of the appropriate locking screwdriver. In task 6, couple participants exhibited an incorrect coupling direction of the Lateral Quick Handle, lacking awareness that the handle should be coupled laterally. Likewise, majorities of the task failure were come from unrecognition of correct tools to use, proper direction to connect, or appropriate order of maneuver. Perhaps, as mentioned above, this might come from lack of attention to, or failure of recalling the user manual and instruction video. Jacob Neilson, in his article about usability heuristics, insisted that better usability designs would let the users recognize rather than recall. That is, the design should minimize the memory loading of the user, rather it should guide the user to use the device in the correct way(Nielsen 1993). Zhang who similarly established 14 principles of usability heuristics in medical devices, also emphasize the same properties (Zhang et al. 2003). For example, one of the suggestions from the interview was it would be better if the handle cannot be combined when it is placed on the wrong side. If so, the users would

recognize the handle must be attached in a certain direction even when they could not recall they needed to.

The most effective strategies in device design to minimize use-related hazards involve optimizing the user interface to be logical and intuitive, facilitating correct user actions while discouraging actions that could lead to harm. Modifying the device design to address use-related hazards is generally more effective than changing labeling or training methods. The consistency of information display and control actions with users' expectations is crucial, aligning with their experiences and likely behaviors. For instance, when the logic of a device departs from users' expectations, such as an electronically driven control dial turning in the opposite direction of previously mechanical dials, the potential for use errors increases (FDA February 3, 2016).

The Surveys and Interviews

The goal of surveys and satisfaction assessment was achieved, as the scores for both were higher than the target score of 3.0. Perhaps, the fact that the testing device is already on the market could introduce a potential bias, leading to the perception of the device and the procedures as safe. User satisfaction survey also met the target standard, above 3.5. The lowest average was 4.4 on questionnaires about readability of the font in the user manual and about clicking sound of torque limiting handle.

The specific comments in the interviews were also valuable. Interviews generate qualitative information regarding the perceptions, opinions, beliefs, and attitudes of individual or groups of device users and patients. Beside the interview comments mentioned above, some comments provided information that are not seen in observational data or the survey scores. For example, some participants mentioned that the sealing of the pouch was somewhat difficult for aseptically opening the implant and the screw. One of them suggested having a separate pouch just for the screw inside the original pouch. There was no use-error associated with this task, and the survey score was 4.2 regarding the safety of the task. In the interviews, users can describe their experiences

with existing devices, specific problems they had while using them, and provide their perspectives on the way a new device should be designed (FDA February 3, 2016).

The Next Step

The expandable cage utilized in the test has already been commercialized, limiting the scope for significant adjustments to the current model. However, the manufacturer can focus on designing instructive materials to mitigate observed errors in tasks. For instance, highlighting distinctions and the sequence of using the two pushers during bone graft insertion can reduce the errors in bone graft insertion. While immediate modifications to the current model may be constrained, the feedback from the usability test serves as valuable input for fundamental enhancements in the next generation of the device. The next generation or subsequent model of the expandable cage would be improved regarding the tasks with least success rate and least satisfactory in the survey.

V. Conclusion

Out of total 24 tasks, 23 tasks achieved the initial test goals of a success rate of 80% or higher and a user error rate of 20% or lower. Five use-error were observed but in critical task which are correctable with minor improvement. In the survey assessing the convenience and safety of task performance in each scenario, the convenience scores averaged above 4.0 points, and the safety scores averaged above 4.2 points. Furthermore, in the satisfaction survey comprising 11 items, all participants' responses surpassed an average score of 4.0 points. The usability test provided useful advice for further improvement of the device. Though unable to perfectly mimic the real-world experiences, the test provided reasonable environment to evaluate the usability of the device with the target user group.

Reference

Baliga, S., K. Treon, and N. J. Craig. 2015. "Low Back Pain: Current Surgical Approaches." *Asian Spine J* 9 (4):645-57. doi: 10.4184/asj.2015.9.4.645.

FDA, U.S. February 3, 2016. Guidance for Industry and Food and Drug Administration Staff: Applying Human Factors and Usability Engineering to Medical Devices. edited by US FDA.

IEC/ISO. 2015. IEC 62366-1 Medical Devices. In *Part 1: Application of Usability Engineering to Medical Devices.* Geneva, Switzerland: International Organization for Standardization.

IEC/ISO. 2019. ISO 14971:2019 Medical devices In *Application of risk management to medical devices*. Application of risk management to medical devices: International Organization for Standardization.

Karsh, B. T. 2004. "Beyond usability: designing effective technology implementation systems to promote patient safety." *Qual Saf Health Care* 13 (5):388-94. doi: 10.1136/qhc.13.5.388.

Katz, J. N., Z. E. Zimmerman, H. Mass, and M. C. Makhni. 2022. "Diagnosis and Management of Lumbar Spinal Stenosis: A Review." *JAMA* 327 (17):1688-1699. doi: 10.1001/jama.2022.5921.

Lewandrowski, K. U., L. Ferrara, and B. Cheng. 2020. "Expandable Interbody Fusion Cages: An Editorial on the Surgeon's Perspective on Recent Technological Advances and Their Biomechanical Implications." *Int J Spine Surg* 14 (s3):S56-S62. doi: 10.14444/7127.

Macki, M., T. Hamilton, Y. W. Haddad, and V. Chang. 2021. "Expandable Cage Technology-Transforaminal, Anterior, and Lateral Lumbar Interbody Fusion." *Oper Neurosurg (Hagerstown)* 21 (Suppl 1):S69-S80. doi: 10.1093/ons/opaa342.

Nielsen, Jakob. 1993. *Usability Engineering*: Academic Press.

Pelayo, S., R. Marcilly, and T. Bellandi. 2021. "Human factors engineering for medical devices: European regulation and current issues." *Int J Qual Health Care* 33 (Supplement_1):31-36. doi: 10.1093/intqhc/mzaa103.

Ravindra, V. M., S. S. Senglaub, A. Rattani, M. C. Dewan, R. Hartl, E. Bisson, K. B. Park, and M. G. Shrime. 2018. "Degenerative Lumbar Spine Disease: Estimating Global Incidence and Worldwide Volume." *Global Spine J* 8 (8):784-794. doi: 10.1177/2192568218770769.

Russ, A. L., and J. J. Saleem. 2018. "Ten factors to consider when developing usability scenarios and tasks for health information technology." *J Biomed Inform* 78:123-133. doi: 10.1016/j.jbi.2018.01.001.

Shin, J., and H. Lee. 2023. "Optimal Usability Test Procedure Generation for Medical Devices." *Healthcare (Basel)* 11 (3). doi: 10.3390/healthcare11030296.

Virzi, R.A. 1992. "Refining the rest phase of usability evaluation: How many subjects is enough?" *Human Factors*:34.

Zhang, J., T. R. Johnson, V. L. Patel, D. L. Paige, and T. Kubose. 2003. "Using usability heuristics to evaluate patient safety of medical devices." *J Biomed Inform* 36 (1-2):23-30. doi: 10.1016/s1532-0464(03)00060-1.

국문 요약

높이 확장형 추간체유합보형재의 사용자적합성 평가

퇴행성요추질환의 유병률 증가에 따라 요추유합술이 활발하게 진행되고 있고 그 기술 또한 발전하고 있다. 본 논문은 새롭게 개발된 높이 확장형 추간체유합보형물의 사용적합성평가에 대한 결과를 탐색해보고자 한다. 높이 확장형 추간체유합보형물은 새롭게 개발된 의료기기로써 기존의 보형물과 사용법의 차이를 두고 있고 때문에 최근 중용되고 있는 사용적합성의 평가가 요구되는 기기이다.

사용적합성평가는 의도된 사용자인 척추수술의 15 명을 대상으로 진행하였고 위험관리 파일에 따라 작성된 24 개의 태스크의 사용시나리오에서 사용자가 태스크를 성공적으로 수행하는가를 관찰하여 평가하였다. 이후 사용자는 각각의 태스크에 대하여 편의성과 안전성에 대한 설문을 진행하였고 의료기기에 대한 총 11 개의 만족도 설문을 실시하였다. 사용자는 인터뷰를 통하여 자유로운 피드백을 제공하였다.

총 24 개의 태스크 중 1 개의 태스크가 80%라는 기준에 미치지 못하는 73%의 성공률을 기록하였고 나머지 태스크는 목표를 충족하며 80% 이상의 성공률을 달성하였다. 각 작업에 대한 편의성 및 안전성에 관한 설문 조사에서 편의성 점수는 4.0 점 이상의 평균을 유지하고 안전성 점수는 4.2 점 이상의 평균을 유지하며 모든 작업은 목표 점수인 3.0 점 이상의 평균 점수를 기록하였다. 총 11 개의 만족도 조사는 4.4 점의 평균을 기록하며 역시 모든 항목에서 3.0 점 이상의 평균 점수를 받았다.

일부 한계에도 불구하고 본 사용적합성평가는 기기의 사용성과 더 나은 개선 방향에 대한 충분한 지표를 제공하였다. 시험된 확장형 추간체유합보형물은 목표 수준의 사용성과 참가자 만족도를 나타냈고 평가로부터 얻게 된 사용자 피드백과

관찰된 사용 오류를 기반으로 의료기기의 지속적인 개선을 이어갈 수 있을 것으로 판단된다.

핵심 되는 말: 높이 확장형 추간체유합보형물, 확장형 추간체보형물, 사용적합성, 사용적합성평가, 의료기기