Ergonomic Reliability Assessment for Operation Process of Medical Device Interface Design Based on OAT and CREAM

Xin Liu, Zheng Liu*, Ziyu Xie and Pengqing Chen
Guangzhou University, No. 230, GuangZhou University City Outer Ring Road, Guangzhou 510006, Guangdong, P.R. China.

E-mail: liu_best@yeah.net

Abstract. This paper uses OAT (Operator Action Tree) and CREAM (Cognitive Reliability and Error Analysis Method) to analyze the interface design of medical equipment according to human error, discuss how to accurately describe the operator's failure behavior, determine the development sequence of important human factors events and the feasible calculation method according to the system response, establish a complete ergonomic reliability assessment modeling by example demonstrations. This study can better describe the relationship between software interface reliability and ergonomic reliability. In addition, this paper makes improvement on the cognitive reliability and error analysis. In this paper, based on the basic concept and core idea of CREAM method, MADM (Multi-attribute Decision-making) is used to improve the weighting factor value of CPCs (Common Performance Conditions) provided by CREAM method. After the OAT model building based on the operation process of medical device interface design, the adjusted failure probability prediction analysis method is applied to obtain the failure probability of OAT basic events, which makes CREAM method more suitable for the medical device accident analysis and risk assessment in complex environment.

Keywords. Ergonomic Reliability, Medical Device, Interface Design

1. Introduction
Ergonomic reliability is a comprehensive discipline that has been developed since the 1950s, which has experienced three generations. The first generation of ergonomic reliability includes the types of human errors, statistical analysis and prediction methods of human error probability based on expert judgment, but lack of psychological basis. Representative methods include THERP (Technic for Human Error Rate Prediction), OAT (Operator Action Tree) etc. The second generation of ergonomic reliability integrates human cognitive psychology and emphasizes the influence of scenarios and environment on ergonomic reliability. The representative methods include CREAM (Cognitive Reliability and Error Analysis Method) and other ergonomic reliability are often used to analyze the operating efficiency and safety of instruments and equipment in high load and dangerous environment. In recent years, there are more researches in the field of medical equipment, and the application of ergonomic reliability in the high-load medical environment is mainly concentrated in the field of clinical surgery. van Rutte et al. identified technical errors and dangerous areas during cannula gastrectomy through OCHRA (Observational Clinical Ergonomic reliability Analysis), so as to improve the operation efficiency and safety of clinical surgery. Sabour et al. used medical
statistics to study the reliability of clinical observational ergonomic reliability and effectively diagnosed the common errors and methods in laparoscopic rectal cancer surgery. Tang et al. [5] proposed a new method of combining objective structure clinical examination with clinical observational ergonomic reliability evaluation to evaluate the surgical skills and cognitive skills of students in the laparoscopic training course. In addition, Tang et al. [6] also reviewed OCHRA and pointed out that increasing the use of OCHRA may improve the prognosis of patients after surgery. The technology should also speed up the clinical absorption of this technology, and the use in conventional surgical practice and surgical training. Davis et al. [7] studied the ergonomic reliability of blood glucose self-monitoring based on portable blood glucose monitor, which has effectively improved the accuracy of blood glucose monitors. Okere et al. [8] developed ultrasonic coupling gel for multiple high infectious diseases in hospitals, which effectively improved the ergonomic reliability of medical equipment in hospitals. There are few researches on the human-machine reliability of medical equipment used in other departments of the hospital. Li Ping et al. [9] analyzed the concept, characteristics and application process of the second generation ergonomic reliability CREAM, and elaborated its application in the analysis of ergonomic reliability of infusion pump, and effectively evaluated the failure probability of operators in the process of using infusion pump. Lin et al. [10] conducted ergonomic reliability of medical equipment based on failure mode and effect analysis and fuzzy semantic theory, and put forward methods that can effectively improve the ergonomic reliability of medical equipment. Ergonomic reliability analysis plays an important role in diagnosing the malpractice of medical equipment design and evaluating the probability of operation error of users [11]. At present, most of ergonomic reliability researches are used in the field of machinery [12], and the existing ergonomic reliability analysis of medical device is mainly used in the hardware products. In addition, further research on ergonomic reliability analysis specifically for the medical device software interface is required.

2. Ergonomic reliability assessment model for the interface design of medical equipment

2.1. Operator action tree model

The OAT is dedicated to operators correctly diagnosing incidents and determining necessary responses in system operations [13]. According to the chronological sequence of the detection, the project uses dichotomy to divide the operation action tree bifurcation. Each bifurcation represents the necessary operation for the medical examiner in virus detection, and there are two possibilities, either success or failure as shown in figure 1. Taking medical device PCR (Polymerase Chain Reaction) as an example, the interface operation of medical device was analyzed as in table 1.

![Operator Action Tree (OAT)](figure1.png)

**Figure 1.** Operator Action Tree (OAT).

**Table 1.** Operator action is decomposed in detail during the interface operation.

| Operation process | Serial number | Operator action in detail |
|-------------------|---------------|--------------------------|
| PCR program       | A1            | Failure of selecting “New experiment”. |
|                   | B1            | Failure of selecting “Detection format”. |
### 2.2. Basic event probability prediction of Operator Action tree

The human error in working environment is largely affected by the psychological cognition and physiological load of human body in the closed state. Based on the second generation of ergonomic reliability theory, by applying cognitive reliability and error analysis theory CREAM can we more comprehensively analyze the probability safety evaluation in the process of operating the PCR detector of virus detection equipment under high workload. The CREAM is a bidirectional human reliability analysis method created by Hollnagel (1998), which allows forward-looking analysis of high-risk systems or processes[14]. CREAM conducted an in-depth study on how PSFs (action formation factors) affect behavior, which is exactly what OAT lacks. At the beginning of the analysis,

| setup and operation | C1 | Failure of setting up “Block size”. |
|---------------------|----|----------------------------------|
|                     | D1 | Failure of setting up “Reaction volume”. |
| Preincubation       | E1 | Failure of setting “Response name” and preincubation cycle number. |
|                     | F1 | Failure of setting up temperature. |
|                     | G1 | Failure of setting up time. |
| Amplification       | H1 | Clicking the add button and failure of setting up the amplification cycles number. |
|                     | I1 | Failure of setting up fluorescence collection function. |
|                     | J1 | Failure of setting up amplification temperature. |
|                     | K1 | Failure of setting up the amplification time. |
|                     | L1 | Failure of setting up acquisition mode of amplification. |
|                     | M1 | Failure of setting up annealing temperature. |
|                     | N1 | Failure of setting up annealing time. |
|                     | O1 | Failure of setting up acquisition mode. |
|                     | P1 | Failure of setting up derived temperature. |
|                     | Q1 | Failure of setting up derived time. |
|                     | R1 | Failure of setting up acquisition mode. |
| Melting curve       | S1 | Failure of setting up cycles number. |
|                     | T1 | Failure of setting up stage 1 temperature. |
|                     | U1 | Failure of setting up stage 1 time. |
|                     | V1 | Failure of setting up a new temperature. |
|                     | W1 | Failure of setting up a new time. |
|                     | X1 | Failure of setting up acquisition mode. |
|                     | Y1 | Failure of setting up a new temperature. |
|                     | Z1 | Failure of setting up acquisition mode. |
| Cooling             | AA1| Failure of setting up cycles number. |
|                     | AB1| Failure of setting up analysis mode. |
|                     | AC1| Failure of setting up temperature. |
|                     | AD1| Failure of setting up time. |
| Sample editor       | AE1| Failure of clicking sample editor and failing to enter the sample edit area. |
|                     | AF1| In selecting workflow part, failure of selecting experiment category. |
|                     | AG1| Failure of selecting sample hole. |
|                     | AH1| In edit Abs quant properties, failure of selecting sample type. |
|                     | AI1| Clicking make replicates, failure of selecting secondary hole. |
|                     | AJ1| Failure of setting up auto Std Curve about standard sample. |
|                     | AK1| Failure of setting up unknown sample. |
|                     | AL1| Failure of setting up subset editing. |
| Results analysis    | AM1| Failure of building new analysis. |
|                     | AN1| In the subset option, Failure of selecting sample distribution area. |
| From A2 to AN2      | The second operator failed to correct the error of the previous operator |
CREAM considered the qualitative and quantitative effects of PSFs at the beginning of the analysis, which is regarded as the reason of behavior. On this basis, CPCs (Common Performance Conditions) were generated. This paper revises the CREAM model to improve the usage and accuracy of CREAM as shown in figure 2.

Figure 2. Flow chart of basic event probability prediction of OAT based on CREAM.

2.2.1. The weight adjustment of CPCs
Since the CPCs weighting factor table provided by CREAM is the average of the weighting factors, it is necessary to adjust the CPC weighting factor value under the actual device application environment to improve the usability. Based on expert researches, MADM (Multi-attribute Decision-making) is applied to calculate the weights of CPCs. After the weights of CPCs are obtained through MADM, multiplying the CPCs weighting factor values to get the modified weighting factor values, which can be used to predict more accurately the probability of ergonomic reliability.

Multi-attribute decision making is used to sort and compare the comprehensive attribute values of factors\(^{15}\). For example, attribute \( b_i \) can make a small difference in attribute values of all schemes, indicating that the impact of attribute \( b_i \) on decision-making and ranking factors is smaller; on the contrary, if attribute \( b_i \) can make a big difference in attribute values of all schemes, the impact of attribute \( b_i \) on decision-making and ranking factors is greater. Therefore, from the point of view of factor ranking, attributes with greater differences in factor attribute values need to be given greater weight values\(^{16}\). As for attribute \( b_i \), the difference between factor \( x_i \) and all other factors is expressed as \( G_{ji}(w) \). \( m \) is the number of the factors about CPCs, \( p \) represents the number of participants as formula (1),

\[
G_{ji}(w) = \sum_{q=1}^{p} \left| b_{ji} w_i - b_{qi} w_i \right| \quad (j \in p, i \in m)
\]
Make,
\[ G_i(w) = \sum_{j=1}^{p} G_{ji}(w) = \sum_{j=1}^{p} \sum_{q=1}^{p} |b_{ji} - b_{qi}| w_i \quad (i \in m) \] (2)

Where, \( G_i(w) \) presents that the total deviation of all schemes for attribute \( b_i \). The selection of weighted vector \( w \) should maximize the total deviation of all attributes for all schemes. Therefore, the construction of the objective function is shown in Formula (3).

\[
\max G(w) = \sum_{i=1}^{m} G_{ji}(w) = \sum_{i=1}^{m} \sum_{j=1}^{p} \sum_{q=1}^{p} |b_{ji} - b_{qi}| w_i
\] (3)

Thus, the weight vector \( w \) obtained is equivalent to the formula (4).

\[
\begin{align*}
\max G(w) &= \sum_{i=1}^{m} \sum_{j=1}^{p} \sum_{q=1}^{p} |b_{ji} - b_{qi}| w_i \\
\text{s.t.} \quad w_i \geq 0, i \in m, \sum_{i=1}^{m} w_i^2 = 1
\end{align*}
\] (4)

Convert equation (4) into Lagrange function,

\[
L(w, \xi) = \sum_{i=1}^{m} \sum_{j=1}^{p} \sum_{q=1}^{p} |b_{ji} - b_{qi}| w_i + \frac{1}{2} \xi \left( \sum_{i=1}^{m} w_i^2 - 1 \right)
\] (5)

The partial derivative is taken and get,

\[
\frac{\partial L}{\partial w_i} = \sum_{j=1}^{p} \sum_{q=1}^{p} |b_{ji} - b_{qi}| + \xi w_i = 0 \quad (i \in m)
\]
\[
\frac{\partial L}{\partial \xi} = \sum_{i=1}^{m} w_i^2 - 1 = 0
\]

Get,

\[
w_i^* = \frac{\sum_{j=1}^{p} \sum_{q=1}^{p} |b_{ji} - b_{qi}|}{\sqrt{\sum_{i=1}^{m} \left( \sum_{j=1}^{p} \sum_{q=1}^{p} |b_{ji} - b_{qi}| \right)^2}}
\] (7)

Normalize \( w_i^* \),

\[
W_i^* = \frac{w_i^*}{\sum_{i=1}^{m} w_i^*} \quad (i \in m)
\] (8)

2.2.2. The adjustment of cognitive failure probability

Four cognitive functions of CPCs provided by CREAM are in table 2. Through the weight factor in table 2, the weighting factor of each CPC for each cognitive activity can be obtained, and then the product of the modified weighting factors of all CPCs under each cognitive activity can be obtained according to formula (9) respectively, that is, the total weighting factor of this cognitive activity can be obtained, then the CFP (Cognitive Failure Probability) of the modified cognitive function failure probability can be deduced by formula (10) according to different circumstance and environment.
very efficient, efficient, inefficient, deficient, OBS, INT, PLAN, EXE

\[
\begin{align*}
    w_a = \frac{w_b \times w_{bcd}}{n \sum_{b=1}^{n} w_b \times w_{bcd} \times \sum_{b=1}^{n} w_{bcd} \times \left( c \in \{ \text{very efficient, efficient, inefficient, deficient} \} \right) d \in \{ \text{OBS, INT, PLAN, EXE} \}}
\end{align*}
\]

(9)

\[
CFP_{\text{modify}} = CFP_{\text{calibration}} \times \prod_{a=1}^{n} w_a (a \in n)
\]

(10)

Where, \( w_a \) is the modified weight of CPCs under various cognitive activities according to the different actual working circumstance, \( w_b \) is the basic weight obtained by MADM, \( w_{bcd} \) is one weighting factor value under a kind of cognitive function.

| Cognitive function | Generic failure type | Basic value |
|--------------------|----------------------|-------------|
| Observation        | O1. Observe the wrong object | 0.001 |
|                    | O2. Identification error | 0.007 |
|                    | O3. Observation not made | 0.007 |
| Interpretation     | I1. False diagnosis | 0.2 |
|                    | I2. Wrong decision | 0.01 |
|                    | I3. Delay of interpretation | 0.01 |
| Planning           | P1. Wrong priority | 0.01 |
|                    | P2. Inadequate planning | 0.01 |
| Execution          | E1. Action of wrong type | 0.003 |
|                    | E2. Action at wrong type | 0.003 |
|                    | E3. Action on wrong object | 0.0005 |
|                    | E4. Action out of sequence | 0.003 |
|                    | E5. Action missed | 0.03 |

2.2.3. Basic event probability prediction

Because the basic failure probability value provided by CREAM is the average failure probability of various events, it is difficult to express the failure probability of operators properly, so it is necessary to modify the basic value of failure probability. According to the 13 failure modes of the four cognitive functions in table 3, the corresponding failure probability in the actual operation is calculated after several tests by professional trained operators. Due to the limited number of tests conducted by medical inspectors in special environment, the final basic value of failure probability is calculated by combining the number of failure modes collected during the test, so that the basic value of failure probability proposed by CREAM can more accurately reflect the actual operation.

3. Case studies of ergonomic reliability assessment in the interaction interface

CPCs weights are analyzed. Through expert researches, 32 valid questionnaires were obtained. From formula (1) to formula (8), the weight vector of experts is obtained as \((0.0410, 0.0076, 0.0293, 0.0205, 0.0564, 0.0264, 0.0528, 0.1073, 0.0440, 0.0258, 0.0317, 0.0674, 0.0070, 0.0275, 0.0328, 0.0358, 0.0199, 0.0252, 0.0123, 0.0322, 0.0176, 0.0457, 0.0416, 0.0234, 0.0387, 0.0135, 0.0252, 0.0129, 0.0211, 0.0275, 0.0170, 0.0129)\). Then, the weight vector is multiplied by expert evaluation matrix, after the normalized treatment, the basis weight vector of CPCs is obtained as \((0.113006, 0.106738, 0.115852, 0.112401, 0.112032, 0.107013, 0.107178, 0.109121, 0.111111, 0.111111, 0.111111)\).
Table 3. Basic weights of CPCs and weighting factors for CPCs\(^{[17]}\)

| CPCs name | Weight  | Level                  | Cognitive function |
|-----------|---------|------------------------|--------------------|
| C1. Proper organization | 0.113006 | Very efficient          | OBS 1 1 0.8 0.8 |
|           |         | Efficient               | INT 1 1 1 1       |
|           |         | Inefficient             | PLAN 1 1 1.2 1.2  |
|           |         | Deficient               | EXE 1 1 2 2       |
| C2. Conditions of working | 0.106738 | Advantageous           | OBS 0.8 0.8 1.0 0.8 |
|           |         | Compatible             | INT 1 1 1 1       |
|           |         | Incompatible           | PLAN 2 2 1 2       |
| C3. Appropriate MMI and operational support | 0.115852 | Supportive            | OBS 0.5 1 1 0.5  |
|           |         | Adequate               | INT 1 1 1 1       |
|           |         | Tolerable              | PLAN 1 1 1 1       |
|           |         | Inappropriate          | EXE 5 1 1 5       |
| C4. Available procedures/plans | 0.112401 | Appropriate           | OBS 0.8 1 0.5 0.8 |
|           |         | Acceptable             | INT 1 1 1 1       |
|           |         | Inappropriate          | PLAN 2 1 5 2       |
| C5. Number of simultaneous goals achieved | 0.112032 | Fewer than capacity   | OBS 1 1 1 1       |
|           |         | Matching current capacity | 1 1 1 1     |
|           |         | More than capacity     | EXE 2 2 5 2       |
| C6. Availability of time | 0.107013 | Adequate               | OBS 0.5 0.5 0.5 0.5 |
|           |         | Temporarily inadequate | INT 1 1 1 1       |
|           |         | Continuously inadequate | PLAN 5 5 5 5       |
| C7. Time of day | 0.107178 | Day-time(adjusted)    | OBS 1 1 1 1       |
|           |         | Night-time(unadjusted) | INT 1.2 1.2 1.2 1.2 |
| C8. Adequate training and preparation | 0.109121 | Adequate, high experience | OBS 0.8 0.5 0.5 0.5 |
|           |         | Adequate, low experience | 1 1 1 1       |
|           |         | Inadequate             | EXE 2 5 5 2       |
| C9. Quality of crew collaboration | 0.116658 | Very efficient         | OBS 0.5 0.5 0.5 0.5 |
|           |         | Efficient               | INT 1 1 1 1       |
|           |         | Inefficient             | PLAN 1 1 1 1       |
|           |         | Deficient               | EXE 2 2 2 2       |

According to the actual situation and environment of the experiment, the actual weights of CPCs are different under three working circumstance. In A working circumstance, proper organization is efficient, conditions of working are compatible, appropriate MMI (Man-machine Interface) and operational support is adequate, availability of procedures is acceptable, number of simultaneous goals is fewer than capacity, availability of time is adequate, time of day is day-time, training and preparation is adequate with low experience, the quality of crew collaboration is efficient. In B working circumstance, every level is same with A working circumstance besides adequate training and preparation which is with high experience. In C working circumstance, every level is same with A working circumstance besides quality of crew collaboration which is deficient. Based on the basic weights, according to formula (9) and (10), the adjustment factor weighting factors for CPCs under the conditions of A, B and C are deduced as shown in figure 3.
Figure 3. Adjustment weighting factors for CPCs under the circumstances of A, B and C

Figure 4. Failure probability of basic event for OAT
The failure probability of basic event about the first operator is shown in figure 4, the failure probability of basic event about the second operator in correcting the first operator is very low, so it is not shown. The failure probability of initial event concerning the second operator is calculated by OAT as follow. In interface design 1, P1=0.1688. In interface design 2, P2=0.0697. So, the ergonomic reliability of interface design 2 is better than that of interface design 1.

4. Conclusion
The reliability of man-machine system depends on the reliability of the people. In the innovative design process, the study of ergonomic reliability of the medical equipment interface design can reduce incidence of medical accidents, the development cost and shorten the development cycle. Therefore, it is very important and practical to establish the ergonomic reliability assessment model of the interactive interface for medical equipment.
(1) OAT divides the process of observation, diagnosis, and response into three stages, but does not consider the role of behavior formation factors. The probability of OAT basic events can be obtained through CREAM. OAT can be combined with CREAM methods, which complements with each other.

(2) The CREAM is one of the representative methods of the second generation of ergonomic reliability analysis. It can trace the root cause and predict the failure probability of human error events. It emphasizes the impact of scenario environment on human cognitive activities and cognitive function errors. Therefore, it can be effectively applied in accident analysis, safety analysis and risk assessment of industrial engineering system. This paper proposes a method for modifying CPCs weighting factors based on MADM, which serves as a reference for accurate prediction of the basic event probability of OAT.

(3) This method can scientifically evaluate ergonomic reliability of the medical equipment interface design, which has strong practicality and is an effective ergonomic reliability assessment method for the interaction interface of medical equipment.

Acknowledgements:
The paper was partially supported by the Youth Innovative Talent Projects from Ordinary University of Guangdong Province (2019WQNCX099), National Natural Science Foundation of China (51905116), Science and Technology Program of Guangzhou, China (201904010463).

References
[1] M. A. B. Alvarenga and P. F. Frutuoso E Melo 2019 A review of the cognitive basis for human reliability analysis. Progress in Nuclear Energy, 117: 103050
[2] J. Mi, Y. Li, W. Peng and et al. 2018 Reliability analysis of complex multi-state system with common cause failure based on evidential networks. Reliability Engineering and System Safety, 174: 71-81
[3] P. van Rutte, S. W. Nienhuis, J. J. Jakimowicz and et al. 2017 Identification of technical errors and hazard zones in sleeve gastrectomy using OCHRA:"OCHRA for sleeve gastrectomy". Surg Endosc, 31(2): 561-6
[4] S. Sabour 2017 Application of objective clinical human reliability analysis (OCHRA) in assessment of technical performance in laparoscopic rectal cancer surgery: common mistakes and methodological issues. Techniques in Coloproctology, 21(5): 401-2
[5] B. Tang, G. B. Hanna, F. Carter and et al. 2006 Competence assessment of laparoscopic operative and cognitive skills: objective structured clinical examination (OSCE) or observational clinical human reliability assessment (OCHRA). World Journal of Surgery, 30(4): 527-34
[6] B. Tang and A. Cuschieri 2020 Objective assessment of surgical operative performance by observational clinical human reliability analysis (OCHRA): a systematic review. Surgical endoscopy, 34(4): 1492-1508
[7] E. Davis and N. S. Latman 2005 Accuracy, reliability, and human factors associated with self-monitoring of blood glucose. Journal of Clinical Engineering, 30(2): 83-90
[8] P. Okere, N. Iloanusi, U. Itanyi and et al. 2019 Low-cost antimicrobial fortification of ultrasound coupling gel: An ergonomic innovation to combat sonology-acquired nosocomial infections. Malawi Med J, 31(1): 45-9
[9] L. I. Ping, W. Zhu, K. Jing-Bo and et al. 2014 Application of Human Reliability Analysis in Risk Assessment of Medical Devices. China Medical Devices, 135(2): 113-115
[10] Q. Lin, D. Wang, W. Lin and et al. 2014 Human reliability assessment for medical devices based on failure mode and effects analysis and fuzzy linguistic theory. Safety Science, 62: 248-56
[11] Y. Li, H. Huang, J. Mi and et al. 2019 Reliability analysis of multi-state systems with common cause failures based on Bayesian network and fuzzy probability. Annals of Operations Research, 1-15
[12] J. Mi, M. Beer, Y. Li and et al. 2020 Reliability and importance analysis of uncertain system with common cause failures based on survival signature. Reliability Engineering & System Safety, 201: 106988
[13] E. Hollnagel 1998 Cognitive reliability and error analysis method. (Horway: Elsevier)
[14] J. L. Deeter 2012 Human reliability analysis in healthcare: application of the cognitive reliability and error analysis method (CREAM) in a hospital setting. (United States: ProQuest LLC)
[15] R. T. D. Prabhakaran, B. J. C. Babu and V. P. Agrawal 2006 Optimum Selection of a Composite Product
System Using MADM Approach. *Materials and Manufacturing Processes*, 21(8): 883-91

[16] X. Ze-Shui 2002 *Study on Methods for Multiple Attribute Decision Making under some Situations*. Nanjing: Southeast University

[17] Y. Kim and J. Park 2019 Incorporating prior knowledge with simulation data to estimate PSF multipliers using Bayesian logistic regression. *Reliability Engineering & System Safety*, 189: 210-7