

# 의료영상관리시스템의 사용성평가 (Usability Test for a Medical Image Filing System)

박재희\*, 이남식\*

## Abstract

In order to provide design concept and guidelines for the user interface of MIDAS™(Medical Image Display and Archiving System), a questionnaire survey and empirical study were conducted. User and task requirements were analyzed based upon survey results. The empirical study was done on the 1.0 version of MIDAS to find out the influence of user characteristics (i.e. job, experiences, etc.) and UI design factors (i.e. layout, wording, procedures) on various usability measures (i.e. performance, satisfaction). To perform empirical tests, eight task scenarios were selected and user interactions were recorded using an auto-logging software. The results show that the doctor group requires more learning time. Also, eight types of user errors such as commission, omission, repeat were identified and the causes of the errors were analyzed related to UI design factors. UI design guidelines were suggested for a new version of medical image filing system.

## I. 서론

개인용컴퓨터의 급속한 보급은 컴퓨터전문가가 아닌 일반인들도 일상생활 속에서 컴퓨터에 자주 노출되게 만들었다. 이에따라 사용하기 복잡하고 어려운 컴퓨터를 어떻게 하면 일반인들도 적은 노력으로 쉽게 배우고, 사용할 수 있도록 하는가에 관한 사용성평가(usability test), 사용자인터페이스(UI; User Interface)설계 등이 HCI(Human Computer Interaction) 분야의 주요 관심사로 대두되게 되었다.

의료영상관리시스템(MIDAS™)은 초음파, 내시경 등의 의료영상자료를 저장, 검색, 전송 할 수 있는 기능들을 가지고 있는 소프트웨어로, 의사들을 사용자집단으로 하고 있다. 그러나 아직 많은 의사들이 컴퓨터 사용 경험이 적은 상태이므로, 사용하기 쉬운 사용자인터페이스의 설계는 의료영상관리시스템 시장에서의 성공에 중요한 요인으로 작용하고 있다고 할 수 있다.

본 연구는 의료영상관리시스템의 사용성(usability) 향상을 위해, 현재의 1.0 판이 안고 있는 UI 설계 상의 많은 문제점을 실증적인 사용성평가 실험을 통하여 파악한 후 향후의 신판 개발을 위한 설계지침을 제공하는데 그 목적을 두었다. 한편 실험의 선행작업으로 사용자특성과 작업내용을 파악하기 위한 설문조사도 수행되었다.

---

\* 한국표준과학연구원 인간공학연구실

- 본 연구는 과학기술처와 (주)메디슨의 공동출연에 의한 특정연구과제비에 의해 수행된 연구임

## II. 사용성평가 실험계획

### 1. 피실험자

피실험자 집단으로 의사군(6명)과 대조집단으로 학생군(6명)을 선정하였다. 의사군과 학생군은 다시 의료영상관리시스템의 사용설명을 읽은 군 3명과 읽지 않은 군 3명으로 구분하였다. 각 군(群) 별 평균 연령과 평균 경력은 다음 Table 1. 과 같다. 의사군과 학생군 간의 컴퓨터 환경에의 노출정도의 차이는 Figure 1. 에서 볼 수 있듯이 의사군은 초보사용자그룹으로, 학생군은 대조집단인 전문사용자 그룹으로 볼 수 있다.

Table 1. 피실험자 군의 구분

피실험자군의 구분		인원	평균연령	평균경력
의사	사용설명서 O	3	31.7 세	전문의과정 2.7 년
	사용설명서 X	3	30.3 세	전문의과정 1.7 년
학생	사용설명서 O	3	25.3 세	박사과정 0.8 년
	사용설명서 X	3	23.7 세	석사과정 1.4 년

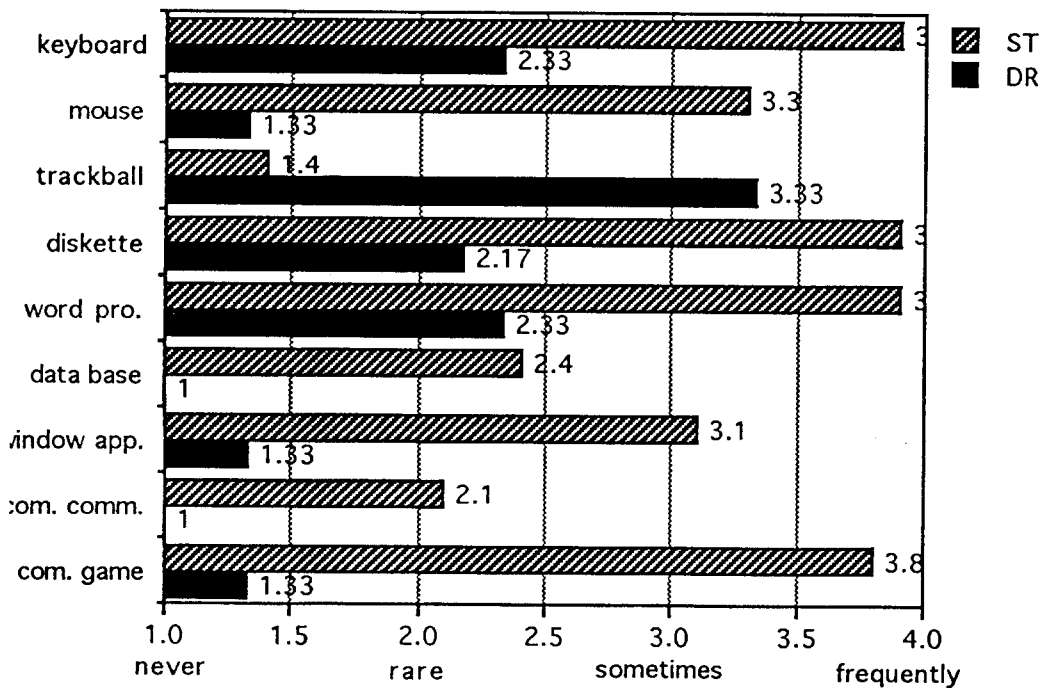


Figure 1. 의사군과 학생군의 컴퓨터환경에의 노출정도

## 2. 실험장비와 소프트웨어

의료영상관리시스템의 사용성평가는 기본적으로 현재 시판되고 있는 MIDAS™의 하드웨어 기본 사양과 동일한 조건하에서 수행되도록 장비를 갖추었으며, 소프트웨어의 경우 MIDAS™ 1.0 한글판 프로그램에, 사용자의 조작내용과 시간을 자동으로 체크하여 출력파일에 그 내용을 기록하는 프로그램(test.exe)과, 조작내용이 담긴파일을 입력으로 받아 이를 다시 재현해 보이는 프로그램 (review.exe)으로 구성되어 있다. 각 프로그램에 대한 설명은 다음과 같다.

- test.exe : 피실험자가 실행하는 프로그램으로, 피실험자의 마우스나 키보드를 이용한 모든 조작을 입력으로 받아 지정된 화일에 그 조작내용과 시간을 텍스트 형식으로 저장하게 된다.
- review.exe : test.exe 에서 생성된 화일을 읽어들이, 피실험자가 수행했던 모든 조작상황을 똑같이 재현하게 되며, 조작시간을 화면의 중앙에 표시하는 화일로 실험자가 분석할 때 사용한다.

## 3. 작업시나리오와 실험방법

사용성평가는 사용자, 작업내용, 사용환경 등에 종속되게 되므로, 정확한 평가를 위해서는 평가의 목적에 맞게 작업내용을 잘 선정해야만 한다. 즉, 소프트웨어의 전반적인 사용성을 평가하기 위해서는 주어진 소프트웨어의 목적과 각 기능들의 사용빈도와 중요성을 감안해 대표작업들을 선정하여야 한다. 대표작업은 몇 개의 이벤트로 구성된 작업 시나리오(scenario)를 가져야 하며 적절한 난이도를 갖도록 하여야 한다. 지나치게 어렵거나, 쉬운 작업에서는 사용성평가 및 개선을 위한 정보를 얻어내기가 어렵다. 이러한 사항들을 고려해 본 연구에서는 의료영상관리시스템의 주요 기능이 고루 포함된 8 개의 작업 시나리오를 도출하였다. 8 개 작업에 포함된 기능들과 작업시나리오의 일례는 Figure 2. 와 같다.

작업 1 : 영상보기 + 화면 Zooming	작업 5 : 음성재생 + 녹음
작업 2 : 모드전환 + 소트 + I버튼의 사용	작업 6 : 거리측정 + 화면처리 + 화면초기화
작업 3 : 디스켓영상언기 + 환자정보입력 + 영상삭제	작업 7 : 문장편집 + 한영전환 + 저장
작업 4 : 문자정보 + 액티브화면 + 영상전송	작업 8 : 영상언기 + 영상저장 + 영상삭제

<p>&lt;작업 1의 시나리오&gt;</p> <p>환자 황지연과 한정수가 있습니다. 두 사람의 영상을 2 개 짜리 화면에 각각 띄운 후에 황지연의 영상을 두 배로 확대해 화면에 크게 띄워줍니다. 황지연 영상의 화면 하단 끝부분에 영문자가 있습니다. 이를 화면에 보이게 한 후 종료합니다.</p>
---

Figure 2. . 각 작업에 포함된 기능과 작업시나리오의 예

실험은 모두 8 개의 작업으로 구성되었으며, 피실험자 당 8 개의 작업을 각각 7 회 반복 수행하도록 하였다. 피실험자는 작업 1 부터 8 까지 수행한 후 다시 반복하는 방법으로 실험을 하였다. 피실험자는 각 실험시작 전 실험의 내용을 충분히 숙지하여 이해하도록 했으며, 실험도중에 시나리오를 계속 참조할 수 있도록 하였다. 사용설명서를 읽는 피실험자 군은 사용설명서를 1 회 읽게 하였고, 실험 전 실험자가 중요사항을 다시 한번 일깨우도록 했다. 사용설명서를 읽지 않는 피실험자 군은 바로 실험에 임하도록 했다. 1 인의 피실험자 당 걸리는 실험시간은 약 2.5~5 시간이 소요되었으며, 휴식은 각 단위 작업 중간중간에 필요할 때 할 수 있도록 하였다. 피실험자가 작업을 수행 못하고 계속적인 에러 상태에 있을 때는, 실험자가 충분한 시간(2분이상)이 경과된 후, 에러를 수정할 수 있는 정보를 주어 작업을 계속 할 수 있도록 하였다. 이 경우 피실험자는 해당 이벤트에서 에러를 발생시킨 것으로 기록되며, 해당 이벤트의 최고제한시간을 적용하도록 하였다.

### III. 실험결과 및 결과분석

#### 1. 사용자군과 사용설명서의 효과

의사군과 학생군 간의 작업 수행수행 시간과 에러 발생을 비교, 분석하기 위하여 분산분석을 실시하였다. 또한 사용설명서를 읽은 군과 안 읽은 군 간의 차이도 알아 보았다. 의사군과 학생군의 작업수행시간에 대해 분산분석을 한 결과 학생군과 의사군 간의 작업수행시간에 차이가 있는 것으로 나타났다 ( $F_{1,671}=319.694, P=0.000$ ), 그러나 사용설명서를 읽은 군과 안 읽은 군 간의 차이는 없는 것으로 나타났다 ( $F_{1,671}=2145, P=0.144$ ). 사용자군-사용설명서, 사용자군-반복회수, 사용자군-작업, 사용설명서-반복회수, 반복회수-작업 간의 교호작용이 있는 것으로 나타났다. 에러를 종속변수로 한 분산분석 결과 도 작업수행시간과 비슷하게 나왔다.

Table 2. 사용자군과 사용설명서의 효과 분산분석

인자	SS	DF	MSE	F	P
주효과	1718083.31	15	114538.89	277.85	0.000
사용자군	131789.01	1	131789.01	319.70	0.000 *
사용설명서	884.38	1	884.378	2.15	0.144
반복회수	576102.73	6	96017.12	232.92	0.000 *
작업	1009307.19	7	144186.74	349.77	0.000 *
2인자교호작용	182881.95	69	2650.46	6.40	0.000
사용자군-설명서	3831.70	1	3831.70	9.43	0.002 *
사용자군-반복	11842.62	6	1973.77	4.79	0.000 *
사용자군-작업	61793.09	7	8827.59	21.41	0.000 *
설명서-반복	7148.69	6	1191.45	2.89	0.009 *
설명서-작업	5384.55	7	769.22	1.87	0.073
반복회수-작업	92881.30	42	2211.46	5.37	0.000 *
설명 잔차	1900965.26 241981.89	84 587	22630.54 412.23	54.90	0.000
계	2142947.15	671	3193.66		

## 2. 학습효과분석

피실험자가 각 작업을 반복 수행할 경우 학습효과가 발생한다. 각 작업의 회수별 수행완료 시간과 에러 수를 파악하여 의료영상관리시스템이 얼마나 배우기 쉬운가 하는 것을 파악해볼 수 있다. 피실험자가 어떤 작업을 7 회 반복 수행할 때, 작업수행시간은 초기에 급격히 줄어들며, 감소율이 서서히 줄어들어 후반부에 갈수록 일정수준에 접근하는 Log 함수 형태를 보여주고 있다(Figure 3. 참조). 작업 수행시간을 반복회수에 대한 Log 함수로 curve fitting 하여 작업수행시간을 예측하는 모형을 수립하였다. 본 실험에서 이 모형의 설명력을 보여주는  $R^2$ 는 0.780~0.987 로그 타당성이 입증되었다. 학습효과 분석을 위한 모형은 다음과 같다.

$$Y = a + b \log(X)$$

Y : 작업수행시간,                      X : 반복회수  
a : 초기 (1회째) 수행시간            b : 학습율, 상수

위의 모형을 적용해 반복몇 회만에 에러없이 정상 속도로 진행하여 표준시간 이내에 도달할 수 있는가를 예측한 결과가 Table 3. 에 나타내었다. 이때, 표준시간은 의료영상관리시스템의 사용에 익숙한 숙련사용자 3 인의 수행완료시간을 측정하여 평균한 값이다. 전체작업의 평균으로 볼 때, 컴퓨터에 많이 노출된 학생군의 경우 5.2 회만에 표준속도에 도달할 수 있는 반면, 초보집단이라고 볼 수 있는 의사군은 적어도 10.6 회를 사용할 때, 표준속도에 이를 수 있는 것으로 나타나고 있음을 알 수 있다. 학생군과 의사군은 사이에는 약 2 배 정도의 수행도 차이가 나타나고 있다. 학생군도 의료용영상관리 시스템에 대해서는 의사군과 마찬가지로 처음 사용하는 것이지만, 다른 일반적 시스템에서 익힌 경험과 지식의 학습 전이가 이루어져 좋은 수행결과를 나타낸 것으로 볼 수 있다.

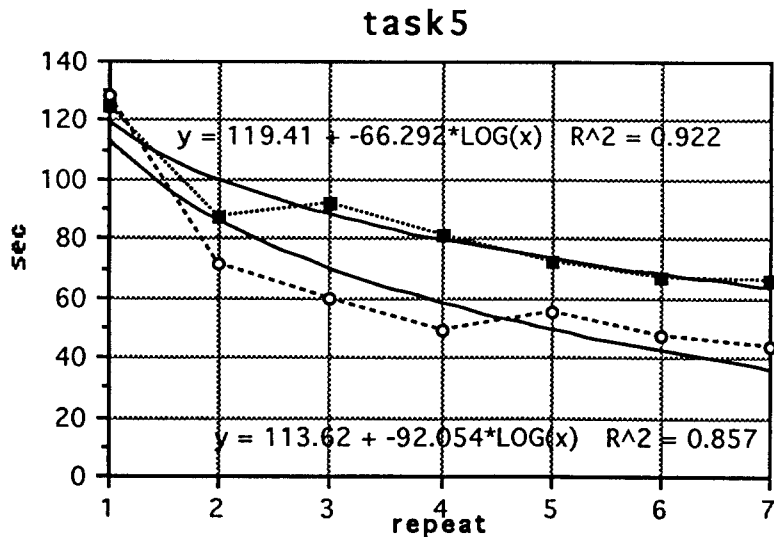


Figure 3. 반복회수와 작업수행시간의 감소

Table 3. 표준시간과 최소요구반복회수

task	Std. T(Y)	DR			ST		
		a	b	repeat(X)	a	b	repeat(X)
1	26.86	89.018	-74.581	6.81	77.587	-71.269	5.15
2	40.79	138.670	-110.870	7.64	94.673	-70.090	5.87
3	76.57	194.350	-134.970	7.46	149.520	-108.740	4.69
4	53.22	155.530	-110.350	8.46	125.610	-106.120	4.81
5	49.76	119.410	-66.292	11.24	113.620	-92.054	4.94
6	44.27	110.760	-73.741	7.97	112.730	-99.954	4.84
7	34.06	118.590	-60.854	24.49	104.460	-84.774	6.77
8	115.12	328.480	-209.400	10.45	247.350	-198.280	4.64
avg				10.56			5.21

### 3. 에러유형과 원인분석

8 개의 작업시나리오 각각은 몇개의 대표가 되는 키이벤트 (key event) 로 나눈 후 키이벤트 전후에서 일어나는 이벤트 조작을 기록 분석해 사용자가 작업을 수행하는 경로(pattern), 패턴( pattern), 수행도중의 에러를 분석해 의료영상관리시스템이 안고 있는 문제점과 원인을 분석하였다. 에러의 유형을 commision, omission, repeat 에러 등 8 개로분류하였고, 이에 따라 1,096 개의 에러를 구분한 결과는 다음 Figure 4. 와 같다. 에러가 발생한 빈도별로키이벤트를 살펴보면 의료영상관리시스템의 어느 부분에서 사용자들이 많은 에러를 발생시키는가를 알 수 있을 것이다. 이러한 분석들을 통하여 현 시스템이 안고 있는 문제점들을 요약하면 다음과 같이 요약할 수 있다.

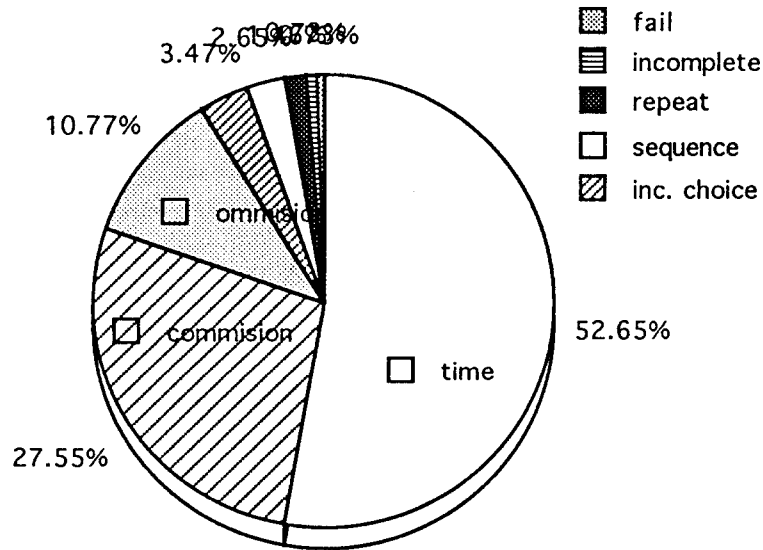


Figure 4. 에러의 유형별 구성

#### o 작업용어의 사용 상의 문제

- 작업용어의 혼돈 : 영상저장(주메뉴)과 영상저장(부메뉴)과 저장 (외부입력장치-> C:)
- 주메뉴와 부메뉴의 동일한 명칭사용 : 영상저장, 영상삭제, 영상저장
- 작업용어의 의미와 기능간의 혼돈 : 영상언기 와 영상받기, I 와 문자, 2X 와 영상처리

#### o 작업진행 상태 정보

- 작업진행정보의 제시미흡 : 작업진행상황에 대한 feed back 정보 부족으로 반복수행 에러 발생
- 작업의 시작과 끝의 불명확성 : 부메뉴작업 후 주메뉴로의 이동 시 상태변환에 대한 차이 불확실
- 액티브화면의 변경 : 액티브 화면 구분이 불명확

#### o 화면, 오브젝트, 메뉴구조의 설계

- 메뉴 구조 설계의 오류 : 상위메뉴와 하위메뉴의 동시전개로 혼란 가중
- 버튼배열의 양립성부족 : 버튼 배열이 인간의 개념양립성과 배치된 경우

### 4. 주관적 사용성 평가결과

사용성평가실험과 별도로 사용자들이 주관적으로 느끼는 사용상의 문제점을 파악하기 위해 의료 영상관리시스템을 충분히 사용한 후 설문을 실시하였다. 설문은 사용성평가실험에 참여했던 12 명의 피 실험자와 예비실험에 참여했던 학생 4 명, 실험에 관계했던 학생 5 명등 의료영상관리시스템에 익숙해 진 총 21 명에게 설문을 실시하였다. 설문을 분석한 결과, 작업상태에 대한 정보의 제공(2.38), 의료행 위와 용어와의 연관(2.42), 전문가를 위한 단축키와 기능키의 사용(2.57) 등에 부정적인 평가를 하였다.

## IV. 결 론

의료영상관리시스템을 4 개의 피실험자군에 대해 사용성평가 실험을 실시하였다. 분석은 크게 실험에 인자 로 채택된 사용자집단(숙련자, 초보자)과 사용설명서에 대한 분산분석, 반복회수에 따른 학습효과분석, 사용자 에러유형과 원인분석 등을 실시하였다. 또한 주관적 사용성평가도 실시하여 의료영상관리시스템이 안고 있는 전반적인 문제점을 파악하였다. 분석결과를 요약하면 다음과 같다.

첫째, 의사집단은 학생집단에 비해 상대적으로 많은 작업수행시간과 에러를 발생시켰다. 이는 소프트웨어 사용성은 그 자체 설계의 적정성 이전에, 사용자집단이 얼마나 컴퓨터 환경에 노출되어 유사한 경험과 지식을 얼마나 가지고 있느냐에 의존하고 있음을 보여주고 있다.

둘째, 반복에 의한 학습효과가 작업수행시간의 단축과 에러발생 감소로 나타났다. 평균적으로 의사군은 10.6 회, 학생군은 5.2 회 반복사용하면, 한 작업을 에러없이 표준속도로 수행할 수 있는 것으로 나타났다.

셋째, 에러가 나타나는 이벤트를 빈도 순으로 파악하여 그 유형을 분석한 결과, 작업용어의 혼돈, 작업상태 정보표시의 미흡, 화면, 오브젝트 설계의 오류 등에 여러 문제가 있는 것으로 나타났다.

넷째, 주관적 사용성평가 결과, 작업상태에 대한 정보의 표시, 용어의 사용, 전문가를 위한 단축키의 사용 등에 낮은 평가가 나타났다.

본 연구에서 평가되고 제시된 내용은 의료영상관리시스템의 추후 사용성 개선에 설계지침으로 반영될 수 있을 것이다. 또한 본 연구에서 사용된 사용성 평가 방법은 다른 소프트웨어에도 일반적으로 적용될 수 있을 것이다. 일반적으로 초보 사용자 집단에게는 특정 소프트웨어의 사용성 보다는 일반 컴퓨터 환경에의 노출정도가 더 중요한 요소로 작용한다는 것을 알 수 있다. 그렇다고 모든 초보자 집단에 다른 기본적 소프트웨어의 사용 경험을 더 요구할 수는 없다. 오히려 지속적인 사용자인터페이스 개선에의 노력을 배가시켜 사용자의 조작에러의 감소와 수행시간을 단축시킬 수 있도록 하는 노력이 더 중요할 것이다.

## 참 고 문 헌

- [ 1 ] Johnson, G.I., Clegg, C.W., and Ravden, S.J., (1989), Towards a practical method of user interface evaluation, Applied ergonomics, Dec., pp.255-260
- [ 2 ] Karat,J., (1988), Software evaluation methodologies in Handbook of human-computer interaction, Elsevier, pp.891-903
- [ 3 ] Marcus, A., (1992), Graphical design for electronic documents and user interfaces, ACM Press
- [ 4 ] Nielsen, J. and Bellcore, (1992), The usability engineering life cycle, IEEE Computer, pp. 12-22
- [ 5 ] Preece, J. and Keller , (1990), Human-computer interaction, prentice Hall
- [ 6 ] Shackel,B. (1989), Human factors and usability , in Human-computer interaction, Prentice Hal, pp.29-41
- [ 7 ] Vizi, R.A., (1992), Refining the test phase of usability evaluation: how many subjectives is enough?, Human Factors, Vol 34(4), pp.457-468
- [ 8 ] Whiteside, J., (1988), Usability engineering: our experience and evolution, in Handbook of human-computer interaction, Elsevier, pp791-817
- [ 9 ] 이남식, 박수찬, 박재희 (1991), 의료용영상전송장치의 User Interface 설계에 관한 연구(1차년도), 한국표준과학연구원
- [10] 이남식, 박재희 (1992), 의료용영상전송장치의 User Interface 설계에 관한 연구(2차년도), 한국표준과학연구원