Korean Journal of Nuclear Medicine Technology, Volume 30, Number 2, Article 22, November 2026.
Korean J Nucl Med Technol 2026;30(2):22. https://doi.org/10.12972/kjnmt.2026.30.2.22
pISSN 1229-9901
eISSN 2982-8406

Original article

문자인식 기반 종사자 건강진단 결과 입력 보조 프로그램

우윤민[1], 이주영[2]
[1]신구대학교 보건의료학부, [2]신구대학교 방사선학과

Optical Character Recognition-Based Data Entry Program for Reporting Radiation Workers' Health Examinations

Yun-Min Woo, Assistant[1] and Joo-Young Lee, Professor[2]
[1]Department of Health & Medical Field, Shingu College
[2]Department of Radiological Science, Shingu College

Corresponding Author: Joo-Young Lee, PhD. Department of Radiological Science, Shingu College. 377 Gwangmyeong-ro, Jungwon-gu, Seongnam, 13174, Republic of Korea. Tel: +82-31-740-1576, E-mail: ljy0706@shingu.ac.kr

Abstract

Purpose: This study developed a data entry program that extracts reporting items from health examination forms of radiation workers processed by optical character recognition, and evaluated its accuracy and work time. **Materials and Methods**: The program extracts 18 required items of the national reporting form and flags uncertain values for staff review. Forms of 100 virtual workers were scanned and processed by optical character recognition. Cell-level accuracy of 16 items read from the forms was compared with an answer key, and work time for 100 workers was estimated with the keystroke-level model. **Results**: Accuracy was 86.0% (95% Confidence Interval, 84.2–87.6%), but only 49.4% in descriptive cells with written content. Error flags identified 61.2–61.6% of errors. With full review, estimated work time decreased from 146.4 to 63.8 minutes (56.4%). **Conclusion**: The program halved work time but should be used with full review by staff.

Keywords: Radiation Workers, Health Examination, Optical Character Recognition, Data Entry, Keystroke-Level Model

서론

현재 원자력관계사업자(이하 사업자)는 원자력안전법(이하 원안법) 시행령 제132조에 따라 방사선작업종사자(이하 종사자) 및 수시 출입자에 대하여 건강진단을 실시하여야 한다[1]. 원안법 시행규칙 제127조와 별표 5에 따라 사업자(신고사용자 제외)는 건강진단 결과를 받은 해당 분기 경과 후 1개월 이내에 원자력안전위원회에 보고하여야 한다[2]. 방사선안전관리자(이하 안전관리자)는 의료기관이 발급한 건강진단표(원안법 시행규칙 별지 제104호의5서식)를 바탕으로 건강진단 결과 보고 서식(방사선방호 등에 관한 기준 별지 제4호의3서식) 및 방사선종사자 방호정보통합관리시스템(Radiation Worker Information System, RAWIS) 업로드 양식을 작성한다[2-4]. 이후 업로드 양식을 RAWIS에 등록하며 건강진단표와 건강진단 결과 보고 서식을 첨부하여 보고한다.

안전관리자는 매년 종사자의 건강진단 결과를 한국원자력안전재단에 보고하며, 이와 관련된 업무량은 종사자 수에 비례한다. 종사자 1인당 보고 양식에는 8개 항목(성명, 주민등록번호, 방사선작업기간, 혈색소량(g/dL), 적혈구수(10^6/μL), 백혈구수(10^3/μL), 혈소판수

Received October 08, 2026 **Revised** October 26, 2026 **Accepted** October 29, 2026

This is an Open Access article distributed under the terms of the Creative Commons CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

(10^3/μL), 특이사항)을, RAWIS 업로드 양식에는 필수 19개 항목(주민등록번호, 성명, 과거질병력, 가족력, 직업력 및 노출력, 방사선 피폭 증상의 유무, 기타 특이사항, 임상진찰, 백혈구수(/μL), 혈소판수(10^3/μL), 혈색소량(g/dL), 적혈구수(10^6/μL), 검진소견, 건강진단서 수령일, 건강진단일자, 병원명, 건강진단의사, 면허번호, 건강진단차수)을 입력해야 한다[3, 4]. 종사자가 100명인 기관이라면 매년 약 2,700개 항목(보고 양식 800개, RAWIS 1,900개)을 안전관리자가 건강진단표를 보며 직접 입력하고 있는 실정이다.

안전관리자가 원본을 보고 수치를 입력하는 과정에서는 필사 오류(Transcription Error)가 불가피하게 발생할 수 있다. Mays 등은 외래 현장검사 혈당 결과 6,930쌍을 분석하여, 수동 입력값의 3.7% (260건)가 장비 연동값과 달랐고 그중 37건(전체의 0.5%)은 20%를 초과하는 차이를 보였다고 보고하였다[5]. 불일치 중 37건은 숫자가 아닌 문자가 그대로 입력된 경우였다[5]. 저자들은 수동 입력이 불가피한 경우의 안전장치로 숫자만 입력할 수 있게 하는 것, 검사항목별로 생리적으로 가능한 범위를 벗어난 값의 입력을 막는 것, 고도의 안전성이 요구되는 경우 같은 값을 두 번 입력하여 대조하는 이중입력을 제안하였다[5].

종이 서식을 자동으로 인식하는 방법도 오류가 발생할 수 있다. Paulsen 등은 환자 설문지 448부를 처리한 연구에서 체크박스를 인식하는 광학마크인식(Optical Mark Recognition, OMR)의 오류율이 수동 이중입력과 차이가 없다고 보고하였다[6]. 반면 환자가 자신의 건강 상태를 0–100의 정수로 적는 EuroQol 시각아날로그척도(EuroQol Visual Analogue Scale, EQ-VAS) 문항을 지능형 문자인식(Intelligent Character Recognition, ICR)으로 인식한 경우의 오류율은 3.53% (9/255)로, 같은 항목의 수동 이중입력(0.78%)보다 유의하게 높았다($P = 0.033$). 이 연구에서는 ICR 결과를 모두 사람이 확인하였는데도 오류가 남았다[6]. 따라서 자동 인식을 도입하려면 정답이 알려진 자료로 정확도를 검증하고, 사람이 확인하는 단계를 절차에 포함해야 한다.

이에 본 연구에서는 광학문자인식(Optical Character Recognition, OCR)을 적용한 건강진단표에서 보고 항목을 추출하고, 불확실한 값을 표시하여 안전관리자가 원본과 대조하도록 돕는 입력 보조 프로그램(이하 입력기)을 개발하였다. RAWIS 등록과 보고는 안전관리자가 수행한다. 본 연구의 목적은 가상 건강진단표로 입력기의 판독 정확도를 평가하고, 안전관리자의 검토를 포함한 작업 시간을 수작업과 비교하는 것이다.

대상 및 방법

1. 연구 설계 및 프로그램 개발

입력기(v1.4.1)는 Python으로 개발하였으며, PDF (Portable Document Format) 파일 읽기에는 pypdfium2, 이미지 속 숫자 인식에는 numpy, 검토 화면에는 tkinter, 엑셀 입출력에는 openpyxl을 사용하였다. 입력기는 OCR 처리된 건강진단표 PDF에서 필요한 정보를 추출하고, 수치·한글 보정과 대상자·기관 목록 대조를 거쳐 RAWIS 필수 19개 항목 중 건강진단차수를 제외한 18개 항목을 요약 양식 엑셀에 기록하였다. 판독이 불확실하거나 보정한 값은 검토 화면에 표시하여 안전관리자가 원본과 대조하게 하였다(Fig. 1). 안전관리자는 검토한 결과를 RAWIS 업로드 양식에 옮기고, 건강진단차수를 직접 입력하도록 하였다. 보고 항목별 처리 방법과 구현 내용을 제시하였다(Table 1, 2).

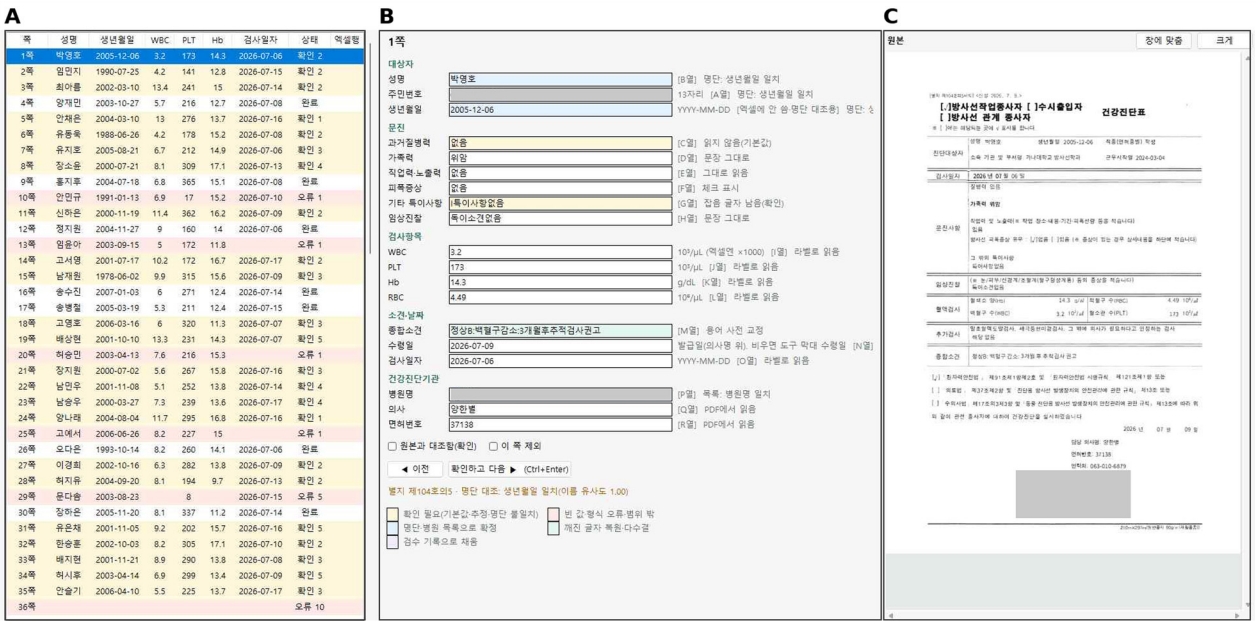


Fig. 1. The review screen of the data entry program (v1.4.1) shows the list of examinees (A), the extracted values with color flags (B), and the original scanned form (C). Red cells indicate errors, yellow cells values requiring confirmation, and blue, cyan, and purple cells values confirmed, restored, or filled by the program (Table 2). All data are virtual; identifiers on the scanned form are masked.

Table 1. Correspondence of health examination form items with the two reporting forms

Item on the health examination form	Report form	RAWIS upload form (column)*	Processing
Name	○	○(B)	Automatic (roster matching)
Date of birth (RRN)	○(RRN)	○(A, RRN)	Automatic (roster matching)
Period of radiation work	○	–	Manual†
Past medical history	–	○(D)	Automatic
Family history	–	○(E)	Automatic
Occupational history and exposure	–	○(F)	Automatic
Presence of radiation exposure symptoms	–	○(G)	Automatic
Other remarks	–	○(H)	Automatic
Clinical examination	–	○(I)	Automatic
Hemoglobin	○(g/dL)	○(L, g/dL)	Automatic
Red blood cell count	○(10^6/μL)	○(M, 10^6/μL)	Automatic
White blood cell count	○(10^3/μL)	○(J, /μL)	Automatic (unit conversion)
Platelet count	○(10^3/μL)	○(K, 10^3/μL)	Automatic
Overall findings (classification)	○(remarks)	○(O)	Automatic
Date of receipt of the examination form	–	○(S)	Automatic (issue date used)
Date of examination	–	○(U)	Automatic
Hospital name	–	○(V)	Automatic (institution list matching)
Examining physician	–	○(W)	Automatic
License number	–	○(X)	Automatic
Examination round	–	○(Y)	Manual†

RAWIS, Radiation Worker Information System; RRN, resident registration number; ○, required item; –, not included.
* ○ indicates the 19 required items of the RAWIS upload form (31 columns)[4]; letters in parentheses are the column labels of the form.
† Not available on the health examination form and entered by the staff (period of work from individual worker records; examination round from previous reports).

Table 2. Requirements and implementation methods of the data entry program

No.	Requirement	Implementation
1	Conversion of scanned forms (image PDF) into text	OCR with Adobe Acrobat Pro; extraction of characters and coordinates from the text layer (pypdfium2); one page processed as one person
2	Extraction of item values	Label-based extraction tolerant of OCR misrecognition; fixed form coordinates used for the examination date, blood test values, and exposure symptom check box when labels are not found
3	Reading of numeric values and conversion to RAWIS format	Separation of values and units; removal of unit digits and restoration of decimal points using format and acceptable ranges*; re-reading of missing values from the page image using digit shapes learned from the same PDF (numpy); WBC ×1,000 (10^3/μL → /μL); dates as YYYYMMDD
4	Correction of broken Korean text (history and findings)	Correction of near misrecognitions and by a term dictionary; restoration of broken standard phrases ("none," "normal") and separated syllables; default values† entered for unreadable or empty fields
5	Distinction of namesakes and verification of identity	Matching with the examinee roster (RRN → date of birth and sex → name similarity)
6	Entry of hospital name, physician, license number, and date of receipt	Hospital name corrected by matching name and telephone number with the institution list‡, otherwise the OCR value used; physician and license number read by labels; issue date used as the date of receipt; unread values filled and hospital and physician names harmonized using values in the same file
7	Error flagging and learning from review	Five cell colors (red, error; yellow, check required; blue, confirmed by roster or list; cyan, restored or corrected; purple, filled from review records); review screen with the original image (tkinter); only confirmed or corrected results learned (history and findings: same correction ≥2 times and ≥80%)
8	Creation of the RAWIS required-item file	Saved to a summary form containing only the required items of the RAWIS upload form (18 columns; openpyxl)
9	Protection of personal information	Local processing without network connection; examinee names, RRNs, and test values not stored in review, learning, or processing records

OCR, optical character recognition; PDF, portable document format; RAWIS, Radiation Worker Information System; RRN, resident registration number; WBC, white blood cell count.
*Ranges for filtering OCR misrecognition, not clinical reference ranges: WBC 2–20 and platelet 80–600 (10^3/μL), hemoglobin 7–20 g/dL, red blood cell 2.5–7 (10^6/μL).
†Normal phrases entered when a field cannot be read; flagged in yellow.
‡531 institutions listed on RAWIS (December 2025)[7].

2. 평가 자료

생성형 인공지능(Artificial Intelligence, AI) 도구인 Claude (Anthropic, San Francisco, USA)를 이용하여 가상 종사자 100명의 건강진단표(1인 1쪽), 대상자 명단, 정답표를 제작하였다. 평가 자료에는 건강진단기관 5곳(이 중 1곳은 기관 목록에 없는 기관), 건강진단의사 10명, 방사선피폭 증상 '있음' 4건, '정상' 외의 내용이 기재된 종합소견 44건(정상A 11건, 정상B 33건)을 포함하였다. 프로그램 개발에 사용한 자료와 평가 자료는 분리하여 분석하였다.

실제 업무에서 안전관리자는 의료기관이 종이로 발급한 건강진단표를 스캔하여 처리하므로, 전자문서로 생성한 가상 건강진단표를 그대로 쓰지 않고 종이 문서 처리 과정을 재현하였다. 생성한 PDF의 글자층을 직접 읽으면 OCR 단계가 생략되어 인식 오류를 평가할 수 없기 때문이다. 건강진단표는 복합기 TASKalfa VFM251ci (Kyocera, Japan)로 컬러, 600 dpi (dots per inch), 표준 품질, A4 설정에서 한 번에 출력한 뒤, 같은 장비로 600 dpi에서 컬러로 스캔하고, Acrobat Pro 2026.002.21931 (Adobe, San Jose, USA)로 OCR(인식 언어: 한국어) 처리하였다. 프린터, 스캐너, OCR 설정은 각각 한 가지로 고정하였고, 인쇄 품질의 영향은 별도로 평가하지 않았다.

본 연구는 사람을 대상으로 하거나 인체유래물·개인정보를 사용하지 않고 가상 자료만 사용하였으므로 기관생명윤리위원회의 심의를 받지 않았다.

3. 정확도 및 실용성 평가

평가용 PDF 100쪽을 입력기를 통해 안전관리자의 검토를 진행하지 않은 상태에서 추출한 결과를 정답표와 칸 단위로 비교하였다(WinPython 3.10.11, pypdfium2 5.13.0, openpyxl 3.0.10). 입력기가 띄어쓰기를 출력하지 않으므로 띄어쓰기를 제외한 글자가 모두 같으면 일치로 판정하였고, 문장부호 차이는 오류로 셌다. 수치는 값, 날짜는 YYYYMMDD 형식으로 비교하였다. 주 평가지표는 건강진단표에서 판독하는 16개 항목(1,600칸)의 정확도로 하고, 대상자 명단에서 가져오는 주민등록번호와 성명을 포함한 18개 항목(1,800칸)의 정확도를 보조지표로 하였다. 정확도는 Wilson 방법의 95% 신뢰구간(Confidence Interval, CI)과 함께 제시하였다[8]. 서술형 7개 항목(과거질병력, 가족력, 직업력 및 노출력, 방사선피폭 증상의 유무, 기타 특이사항, 임상진찰, 종합소견)은 문자 오류율(Character Error Rate, CER; 편집거리 ÷ 정답 글자 수)을 구하고, 모든 칸에 기본값만 입력한 경우와 비교하였다[9, 10]. 또한 판독하지 못한 칸에 입력기가 넣는 정상 문구('없음', '정상' 등, 이하 기본값)만으로 모든 서술형 칸을 채운 경우의 정확도를 구하여 비교 기준으로 삼았다. 오류 표시 기능은 표시 민감도(표시된 오류 ÷ 전체 오류)와 정밀도(표시된 오류 ÷ 표시된 칸)로 평가하였다[11].

작업 시간은 100명의 필수 19개 항목을 RAWIS 업로드 양식에 오류 없이 채우는 과제로 정의하고, 키 입력 수준 모델(Keystroke-Level Model, KLM)로 추정하였다(Table 3)[12, 13]. 수작업은 칸마다 원본 읽기, 두벌식 타수만큼의 입력, 다음 칸 이동으로 계산하였다. 입력기는 모든 칸을 원본과 대조하고 틀린 칸은 재입력과 부분 수정 중 빠른 방법으로 교정하는 경우를 주 분석으로 하였다. 보조 분석으로 틀린 칸을 모두 재입력하는 경우, 오류 위치를 안다고 가정하고 쪽 확인만 하는 경우, 표시된 칸만 확인하는 경우를 함께 산출하였다. 민감도 분석에서는 타자 속도(K 0.12, 0.20, 0.28, 0.50초)와 심리적 준비(Mental Preparation) 시간(M 1.20, 1.35초)의 값을 바꾼 8가지 조합을 두 조건에 같게 적용하였다(Table 3). 손익분기 오류율은 입력기 조건이 수작업과 같은 시간이 되는 항목별 오류율로, (수작업 1칸 시간 − M) ÷ (교정 1칸 시간 − M)으로 구하였다.

Table 3. Operators and values of the keystroke-level model

Operator	Description	Time (s)
K	Keystroke (average typist, 40 words per minute)	0.28
P	Pointing to a target with the mouse	1.10
B	Pressing or releasing the mouse button (click = BB)	0.10
H	Moving the hand between keyboard and mouse	0.40
M	Mental preparation*	1.35
R	System response	0

* Locating, reading, comparing, and judging values in the original form.
Values are from Card et al. [12], except B and M = 1.20 s, which are from Kieras [13]. R was set to 0 to include only user operations.

결과

1. 판독 정확도

16개 항목의 칸 정확도는 86.0% (1,376/1,600; 95% CI 84.2–87.6%)였고, 18개 항목은 87.0% (1,566/1,800; 95% CI 85.4–88.5%)였다(Table 4). 방사선피폭 증상의 유무가 98.0%로 가장 높고, 기타 특이사항(56.0%)과 종합소견(70.0%)이 가장 낮았다.

서술형 7개 항목의 정확도는 80.0% (560/700)로, 모든 칸에 기본값만 입력한 경우(78.0%, 546/700)보다 2.0%p 높았으며, CER은 17.0%였다. 기재 내용이 있는 칸의 정확도는 49.4% (76/154)였다. 병원명·건강진단의사·면허번호를 제외한 13개 항목의 오류 193칸은 문자 오인식(49.7%), 공란(21.8%), 기재 내용 누락(12.4%) 순으로 많았다. 쪽 전체의 양식 인식에 실패한 경우는 5쪽(36, 58, 76, 78, 97쪽)이었다.

2. 오류 표시 성능

16개 항목의 표시 민감도는 61.2–61.6% (137–138/224), 18개 항목은 62.8–63.2% (147–148/234)였다. 18개 항목의 정밀도는 49.5% (147/297)였다. 표시되지 않은 오류는 86–87칸이었다. 방사선피폭 증상 '있음' 4건은 모두 '있음'으로 판독되었으나, 그 중 2건은 괄호 안의 증상 내용이 누락되었다. 수령일 오류 5칸은 모두 평가 자료에서 가장 다수의 발급일인 20260713으로 입력되었다. 두 항목의 오류 7칸은 하나도 표시되지 않았다(Table 4).

3. 작업 시간

100명의 작업 시간은 수작업 146.4분, 입력기(주 분석) 63.8분으로 56.4% 줄었고, 민감도 분석에서는 36.1–69.6%였다(Table 5). 1회만 하는 조작을 제외한 1인당 소요시간은 수작업 87.7±8.4초, 입력기 37.4±13.3초였다. 100명 중 98명은 입력기가 수작업보다 빨랐고, 느렸던 2명(58, 78쪽)은 모두 양식 인식에 실패한 쪽이었다. 입력기 조건 시간의 87.8%는 원본 대조와 교정이었다. 손익분기 오류율을 넘은 항목은 없었으며, 오류율이 가장 높은 기타 특이사항(44.0%)은 모든 칸이 틀려도 입력기가 수작업보다 빨랐다.

Table 4. Reading accuracy and error flag sensitivity by item (n = 100)

Item	Accuracy, % (n/N)	95% CI	Flag sensitivity, n/N
Items read from the form			
Past medical history	82.0 (82/100)	73.3–88.3	6/18
Family history	82.0 (82/100)	73.3–88.3	6/18
Occupational history and exposure	89.0 (89/100)	81.4–93.7	8/11
Presence of radiation exposure symptoms	98.0 (98/100)	93.0–99.4	0/2
Other remarks	56.0 (56/100)	46.2–65.3	29/44
Clinical examination	83.0 (83/100)	74.5–89.1	8/17
White blood cell count	92.0 (92/100)	85.0–95.9	8/8*
Platelet count	91.0 (91/100)	83.8–95.2	9/9*
Hemoglobin	92.0 (92/100)	85.0–95.9	8/8*
Red blood cell count	92.0 (92/100)	85.0–95.9	8/8*
Overall findings	70.0 (70/100)	60.4–78.1	15/30
Date of receipt of the examination form	95.0 (95/100)	88.8–97.8	0/5
Date of examination	85.0 (85/100)	76.7–90.7	14/15*
Hospital name	92.0 (92/100)	85.0–95.9	8/8
Examining physician	84.0 (84/100)	75.6–89.9	5/16
License number	93.0 (93/100)	86.3–96.6	5/7
Subtotal (16 items)	86.0 (1,376/1,600)	84.2–87.6	137/224
Items matched with the roster			
Resident registration number	95.0 (95/100)	88.8–97.8	5/5
Name	95.0 (95/100)	88.8–97.8	5/5
Total (18 items)	87.0 (1,566/1,800)	85.4–88.5	147/234

CI, confidence interval

Accuracy was judged by cell-level agreement ignoring spaces, with punctuation differences counted as errors; confidence intervals were calculated by the Wilson method. Flag sensitivity is the number of error cells flagged in red or yellow divided by the number of error cells, restored from the processing records. Descriptive items were past medical history, family history, occupational history and exposure, presence of radiation exposure symptoms, other remarks, clinical examination, and overall findings.

* Lower bound, because correction and image-reading flags were not stored in the processing records.

Table 5. Estimated work time for manual entry and program use (keystroke-level model, 100 examinees)

Condition	Total time (min)	Per person (s)*	Reduction (%)	Remaining errors (cells)†
Manual entry	146.4	87.8	–	0
Program, full review, optimal correction (main analysis)	63.8	38.3	56.4	0
Program, full review, re-entry of all errors	75.3	45.2	48.5	0
Program, page confirmation only‡	28.6	17.1	80.5	0
Program, review of flagged cells only§	24.6–34.8	14.8–20.9	76.2–83.2	86–87

K = 0.28 s and M = 1.35 s. Reduction = (manual time − program time) ÷ manual time.
* Total time ÷ 100; means excluding one-time operations are given in the text.
† Assumed by the model; entry and review errors by staff were not included.
‡ Theoretical lower limit assuming that the reviewer already knows the error locations.
§ Ranges reflect the lower and upper bounds of flagged cells; unflagged errors remain, so the final result differs from manual entry.

고찰

입력기는 16개 항목의 86.0%를 정확히 판독하였고, 모든 칸을 검토하여도 작업 시간을 56.4% 줄였다. 입력기 조건 시간의 87.8%가 원본 대조와 교정이었으므로, 작업 시간이 단축된 것은 값을 직접 입력하던 일이 입력된 값을 원본과 대조하는 일로 바뀌었기 때문이다.

수작업 입력에서도 오류가 발생할 수 있으며, Mays 등에 의하면 혈당 수치의 수동 전사(Manual Transcription)에서는 3.7%의 오류가 보고되었다[5]. 그러나 본 연구의 KLM은 수작업과 입력기 조건 모두 안전관리자가 실수하지 않는다고 가정하였으므로, 오류 측면의 차이는 반영하지 못하였다. 두 방법은 오류가 생기는 단계도 다르다. 수작업의 오류는 입력 과정에서 생기며 따로 표시되지 않지만, 입력기의 오류는 판독 단계에서 생기며 일부가 표시되어 검토에서 찾을 수 있다. 입력기는 수치 항목, 주민등록번호, 성명, 병원명의 오류를 모두 표시하였으므로, 이들 항목에서는 수작업보다 오류를 줄일 가능성이 있다. 반면 서술형 항목은 오류의 상당수가 표시되지 않아 이러한 효과를 기대하기 어렵다. 다만 실제로 오류가 줄어드는지는 안전관리자가 입력기를 직접 사용하는 조건에서 측정해야 한다.

정확도는 항목의 형식에 따라 차이가 있었다. 체크 표시로 고르는 방사선피폭 증상의 유무는 98.0%로, 체크 표시는 모두 맞게 판독되었다. 수치 항목은 91.0–92.0%였으나, 서술형 항목은 80.0%로 모든 칸에 기본값만 넣은 경우(78.0%)와 거의 같았고, 기재 내용이 있는 칸은 49.4%만 정확하였다. Paulsen 등도 체크박스는 수동 이중입력만큼 정확하였으나 손글씨 숫자는 오류가 더 많았다고 보고하였으며, 본 연구에서도 기재 형식이 자유로운 항목일수록 오류가 많았다[6]. 13개 항목 오류의 49.7%는 유사 글자 오인식이었으며, OCR 정확도는 언어와 이미지 상태에 따라 크게 달라지는 것을 확인할 수 있다[10].

오류 표시 민감도는 약 61%였다. 표시되지 않은 오류의 약 80%는 서술형 항목에 있었으며, 글자 일부만 틀려 올바른 문구처럼 보이는 값이었다(예: '간경화'를 '간경화'로 판독). 이러한 값과 다수의 발급일로 잘못 입력된 수령일은 빈 값·형식·범위 검사를 모두 통과하므로, 표시가 없어도 원본과 대조해야 한다[5]. 표시된 칸만 확인하면 표시에 의존하는 자동화 편향(Automation Bias)으로 오류를 놓칠 수 있으므로, 모든 칸을 검토해야 한다[14]. 그러나 Barchard 등은 사람이 원본과 눈으로 대조하는 방법이 이중입력보다 오류가 많았다고 보고하였다[15]. 따라서 검토 후 오류가 남지 않는다는 주 분석의 가정은 실제 작업에서는 성립하지 않을 수 있다.

본 연구는 생성형 AI로 만든 가상 건강진단표로 평가하였고, 5개 기관으로 제한한 건강진단표를 프린터, 스캐너, OCR 프로그램 각 1종의 한 가지 설정으로만 처리하였다. 인쇄 품질과 스캔 설정은 OCR 결과에 영향을 미치므로 다른 설정에서는 정확도가 달라질 수 있다[10]. 건강진단표 서식은 법령으로 정해져 있으나, RAWIS 목록에 수록된 건강진단기관은 531곳이며 기관마다 기재 방식과 인쇄 상태가 달라 실제 건강진단표에서는 입력기의 성능이 달라질 수 있다[7]. 개발자가 직접 입력기의 성능을 평가하였으며, 개발 자료와 평가 자료는 분리하였으나 같은 방식으로 만든 가상 자료였다. KLM은 오류 없는 숙련자를 가정하며 예측 오차가 약 21%이다[12]. 향후 실제 건강진단표와 안전관리자를 대상으로 작업 시간과 최종 오류를 측정할 필요가 있다.

결론

OCR 처리한 종사자 건강진단표에서 RAWIS 필수 항목을 추출하는 입력기를 개발하여 가상 자료로 평가하였다. 서술형 항목 중 기재 내용이 있는 칸의 정확도는 49.4%였고, 전체 오류의 38.4–38.8%가 표시되지 않는 한계가 있었다. 그러나 입력기의 16개 항목 판독 정확도는 86.0%였으며, 모든 칸을 검토하는 경우에도 100명의 작업 시간이 146.4분에서 63.8분으로 56.4% 단축되는 것으로 추정되었다. 이러한 결과는 입력기가 안전관리자의 전체 검토를 전제로 입력 시간을 줄이는 보조 도구로 활용될 수 있음을 시사한다.

References

1. Ministry of Government Legislation. Enforcement Decree of the Nuclear Safety Act. Presidential Decree No. 36664 (September 8, 2026). Korea Law Information Center. Available at: https://www.law.go.kr. Accessed September 16, 2026.
2. Ministry of Government Legislation. Enforcement Rule of the Nuclear Safety Act. Ordinance of the Prime Minister No. 2154 (September 8, 2026). Korea Law Information Center. Available at: https://www.law.go.kr. Accessed September 16, 2026.
3. Nuclear Safety and Security Commission. Standards for radiation protection, etc. NSSC Notice No. 2026-12 (July 9, 2026). Korea Law Information Center. Available at: https://www.law.go.kr. Accessed September 16, 2026.
4. Korea Foundation of Nuclear Safety. Notice on RAWIS reporting following amendments to regulations on health examinations of radiation workers, including the Enforcement Rule of the Nuclear Safety Act (July 10, 2026). Attachment 3: New health examination registration form (as of July 10, 2026) [xlsx]. Radiation Worker Information System. Available at: https://nsis.kofons.or.kr/mng/home/rawis/meet/ntc. Accessed September 16, 2026. (Korean)
5. Mays JA, Mathias PC. Measuring the rate of manual transcription error in outpatient point-of-care testing. J Am Med Inform Assoc. 2019;26(3):269-72. DOI: https://doi.org/10.1093/jamia/ocy170.
6. Paulsen A, Harboe K, Dalen I. Data entry quality of double data entry vs automated form processing technologies: a cohort study validation of optical mark recognition and intelligent character recognition in a clinical setting. Health Sci Rep. 2020;3(4):e210. DOI: https://doi.org/10.1002/hsr2.210.
7. Korea Foundation of Nuclear Safety. Provision of data on health examination institutions (December 1, 2025). Attachment: List of health examination institutions [xlsx]. Radiation Worker Information System. Available at: https://nsis.kofons.or.kr/mng/home/rawis/meet/ntc. Accessed September 16, 2026. (Korean)
8. Wilson EB. Probable inference, the law of succession, and statistical inference. J Am Stat Assoc. 1927;22(158):209-12. DOI: https://doi.org/10.1080/01621459.1927.10502953.
9. Navarro G. A guided tour to approximate string matching. ACM Comput Surv. 2001;33(1):31-88. DOI: https://doi.org/10.1145/375360.375365.
10. Hegghammer T. OCR with Tesseract, Amazon Textract, and Google Document AI: a benchmarking experiment. J Comput Soc Sci. 2022;5(1):861-82. DOI: https://doi.org/10.1007/s42001-021-00149-1.
11. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Inf Process Manag. 2009;45(4):427-37. DOI: https://doi.org/10.1016/j.ipm.2009.03.002.
12. Card SK, Moran TP, Newell A. The keystroke-level model for user performance time with interactive systems. Commun ACM. 1980;23(7):396-410. DOI: https://doi.org/10.1145/358886.358895.
13. Kieras D. Using the keystroke-level model to estimate execution times (1993; modified by A. Hornof, April 19, 2001). University of Michigan. Semantic Scholar. Available at: https://www.semanticscholar.org/paper/Using-the-Keystroke-Level-Model-to-Estimate-Times-Kieras/9124d25a381ef77fd8f4ec326aa9679a7b6853d3. Accessed September 30, 2026.
14. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc. 2012;19(1):121-7. DOI: https://doi.org/10.1136/amiajnl-2011-000089.
15. Barchard KA, Freeman AJ, Ochoa E, Stephens AK. Comparing the accuracy and speed of four data-checking methods. Behav Res Methods. 2020;52(1):97-115. DOI: https://doi.org/10.3758/s13428-019-01207-3.