Revista Española
de Educación Médica

Workplace-Based Assessment of Admission Documentation and Medical Orders in Specialty Training: Development and Initial Validation of a Multidomain Tool.

Maryam Hosseini[1], Mahvash Alizade Naini[2], Rita Rezaee[3], Masoumeh Hosseini[4], Sepideh Jahandideh[5].

[1] Shiraz University of Medical Sciences, Shiraz, Iran, Maryam.hosseini6072@gmail.com
[2] Department of Internal Medicine, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran, mahvashnaini@gmail.com
[3] Clinical Education Research Centre, Shiraz University of Medical Sciences, Shiraz, Iran, ritarezaee@sums.ac.ir
[4] Hafez Hospital, Shiraz University of Medical Sciences, Shiraz, Iran, masoumehhossini57@gmail.com
[5] Faculty of Health, Southern Cross University, Gold Coast, Australia, sepideh.jahandideh@scu.edu.au
* Correspondencia: sepideh.jahandideh@scu.edu.au

Received: 6/23/26; Accepted: 7/23/26; Published: 7/27/26

Abstract.

Admission documentation and medical orders are important workplace activities in specialty training because they communicate clinical reasoning, guide patient care, and provide evidence of professional competence. However, structured tools for assessing these practices in postgraduate clinical settings remain limited. This study aimed to develop a multidomain assessment tool and collect initial validity evidence regarding its content validity, internal structure, and internal consistency for assessing the completeness of admission documentation and medical orders completed by doctors in specialty training. **Method**: This cross-sectional methodological study was conducted in three phases. Candidate domains and items were identified through a structured literature review and expert discussion, yielding a preliminary five-domain tool comprising 20 items. Content validity evidence was examined using a two-round Delphi process with 13 internal medicine faculty members. Evidence based on internal structure was assessed using confirmatory factor analysis with data from 100 doctors in specialty training in internal medicine, cardiology, and neurology. Internal consistency was evaluated using Cronbach's alpha. The finalized tool was then applied to 200 inpatient medical records from six teaching hospitals. **Results:** Expert consensus supported the relevance and clarity of all five conceptually defined domains and 20 items. However, CFA suggested a four-factor empirical structure, with clinical investigations and medication orders combined into a single factor. Model fit was acceptable based on RMSEA = 0.063, CFI = 0.920, and SRMR = 0.081, although the chi-square test was statistically significant ($p < 0.001$). Internal consistency was acceptable (Cronbach's alpha = 0.79). Application of the tool to 200 records identified variation in documentation completeness across domains and specialties, with internal medicine trainees achieving the highest overall scores. All domain scores were significantly correlated with the total score ($p < 0.001$). **Discussion**: The findings provide initial validity evidence supporting the use of the tool for structured assessment of admission documentation and medical orders in specialty training. At this stage, the tool appears most appropriate for formative feedback and identification of documentation gaps. Further studies are needed to examine inter-rater reliability, response processes, relationships with other assessment measures, and the educational consequences of using the tool in workplace-based assessment programs.

Keywords: workplace-based assessment, postgraduate medical education, specialty training, clinical documentation, medical orders, formative feedback, clinical competence, patient safety.

1. Introduction

Ensuring patient safety and high-quality clinical care is a core responsibility of healthcare systems and postgraduate medical education programs worldwide. Specialty training plays a central role in preparing doctors to make sound clinical decisions, communicate them clearly, and work effectively within interprofessional teams. Contemporary competency-based medical education frameworks therefore emphasise the need for valid, reliable, workplace-based assessment tools that capture observable clinical competence in authentic settings and support both trainee learning and patient safety outcomes (1-3).

Clinical documentation is a core component of professional competence and a key means through which doctors communicate clinical reasoning, decisions, and care plans. High-quality documentation supports continuity of care, interdisciplinary collaboration, quality improvement, research, and medico-legal accountability (4). In contrast, omissions, ambiguity, and lack of standardisation are associated with medical errors, delays in care, and preventable adverse events, making documentation quality an important determinant of patient safety and a relevant focus for clinical education (4, 5). Within medical records, admission documentation and medical orders are particularly high-impact because they capture the initial clinical picture and translate clinical reasoning into actionable instructions. These entries direct medications, investigations, monitoring, diet, activity, and other therapeutic or supportive interventions delivered by the broader healthcare team. Errors or omissions may lead to inappropriate care, miscommunication, and patient harm. Ordering errors, including incorrect doses and wrong medication selection, have been associated with severe harm and, in some cases, death (6). Accurate documentation, therefore, requires sound clinical judgment together with adherence to legal, institutional, and professional standards.

From an educational perspective, completing admission documentation and medical orders is an integrative workplace activity that reflects clinical reasoning, prioritisation, synthesis of patient information, and attention to safe care processes. Clinical documentation is not only a record of care but also a visible product of trainees' thinking. It therefore offers a valuable, yet underused, opportunity for assessment and feedback in postgraduate medical education (7, 8). Despite this potential, feedback on documentation and order writing is often informal, inconsistent, or absent due to time constraints, the lack of standardised criteria, and the limited availability of validated assessment tools. Several instruments have been developed to assess clinical documentation quality, including QNOTE, PDQI-9, IDEA, and CRANAPL, which primarily focus on narrative clinical notes [9-12]. Recent scoping work indicates that documentation assessment tools are diverse; however, few are designed to evaluate the broader set of admission documentation and medical orders relevant to postgraduate specialty training, and many lack evidence about applicability across learner levels and clinical settings (13). Against this background, the present study aimed to develop a structured, multidomain tool and collect initial validity evidence for assessing the completeness of admission documentation and medical orders among trainees in internal medicine, cardiology, and neurology. By defining key domains through expert consensus and examining the tool's validity and reliability, this study sought to support formative and summative assessment, strengthen documentation-related feedback, and improve trainee learning, clinical documentation practice, and patient safety.

According to Messick's unified validity framework, validation is an ongoing process that requires multiple complementary sources of evidence, including content, response processes, internal structure, relations with other variables, and consequences of testing. Because instrument development generally occurs in stages, initial studies commonly focus on content validity and internal structure before evaluating additional sources of validity evidence (14). Accordingly, the present study aimed to develop the instrument and collect initial evidence regarding its content validity and internal structure.

2. Methods

2.1 Study Design

This cross-sectional methodological study aimed to develop and validate a tool for assessing the completeness of admission documentation and medical orders completed by doctors in specialty training. The study was conducted in three sequential phases: (1) identification of candidate domains and items, (2) content validation using expert consensus, and (3) assessment of internal structure, reliability, and application to documentation performance.

2.2 Participants and Setting

The study was conducted across six teaching hospitals affiliated with Shiraz University of Medical Sciences (Shahid Faghihi, Namazi, Chamran, Zeynabieh, Dashtgheib, and Khalili). Participants differed by study phase. For content validation, 13 internal medicine faculty members were recruited from six Iranian medical universities: Shiraz, Tehran, Gilan, Isfahan, Hormozgan, and Bushehr. Eligibility criteria for expert participation included specialty training in internal medicine, current academic faculty appointment, and more than 5 years of experience in teaching and research. Experts were selected using non-probability purposive sampling, supplemented by snowball sampling. For construct validation, 100 doctors in specialty training in internal medicine, cardiology, and neurology participated. These three specialties were included in the psychometric testing phase because they represented the largest available groups of eligible trainees who routinely completed admission documentation and medical orders in inpatient settings. In the application phase, 200 inpatient medical records from the six teaching hospitals were reviewed. For the application phase, inpatient medical records were selected using systematic random sampling from computerized hospital records. After generating an ordered list of eligible admissions for the study period, one record was selected at a fixed interval of every 300 files. Eligibility was limited to records containing admission documentation and medical orders completed by doctors in specialty training, and sampling was undertaken to ensure representation across specialties and training levels.

2.3 Development of the Evaluation Tool

A structured literature review was undertaken to identify relevant documentation standards, candidate domains, and assessment criteria for development of the preliminary tool. Literature published between 2013 and 2023 was searched in PubMed, Scopus, Google Scholar, and ERIC (Education Resources Information Center) using combinations of terms related to admission documentation, clinical documentation, medical orders, physician orders, order writing, documentation quality, completeness, assessment, medical education, trainees, residents, and specialty training. This date range was selected to capture contemporary literature relevant to current documentation practices and assessment approaches. Titles and abstracts were screened for relevance, followed by full-text review of potentially eligible sources. Of the 1,345 records retrieved, 44 were judged relevant and were used to inform the preliminary framework.

Findings from the literature review were synthesized through structured discussions with the research team and faculty experts in internal medicine. These discussions involved the research team and faculty experts in internal medicine and focused on refining domains, clarifying item definitions, and resolving discrepancies identified during the literature review. This process produced five conceptually defined domains and 20 items: legal components, initial patient assessment, care orders, clinical investigations, and medication orders.

Domain 2 was retained because, in the local admission documentation format evaluated in this study, initial patient assessment elements were recorded alongside medical orders and were considered essential to the completeness and clinical usability of admission documentation.

- Legal components: legibility, signature, date/time, and ward

- Initial patient assessment: diagnosis, clinical status, activity level, diet, and allergies
- Care orders: vital signs, IV fluids, urgent notifications, and nursing instructions
- Clinical investigations: laboratory, radiologic, and paraclinical studies
- Medication orders: essential and conditional medications

2.4 Content Validation

Content validity was assessed using a two-round Delphi process with 13 internal medicine specialists from leading Iranian universities. Experts were recruited using purposive and snowball sampling. In the first Delphi round, panellists rated the relevance and appropriateness of each item across the five domains using a structured checklist. Eighteen items were accepted in the first round, and two were revised and reassessed in the second round. The level of agreement among experts was quantified using Kendall's coefficient of concordance, and consensus was achieved upon completion of the second round.

2.5 Construct Validity and Reliability

Evidence based on internal structure was examined using confirmatory factor analysis (CFA). A sample of 100 doctors in specialty training was used for the 20-item tool, yielding 5 participants per item. Although this ratio has been used as a minimum practical guideline for preliminary CFA, the sample size was modest for a multidomain model; therefore, the CFA findings were interpreted as initial evidence rather than definitive confirmation of the factor structure. The finalized form was administered to doctors in specialty training in internal medicine, cardiology, and neurology. Data were analyzed using LISREL version 8.80. Model fit was assessed using the available CFA outputs, including the chi-square test, root mean square error of approximation (RMSEA), comparative fit index (CFI), and standardized root mean square residual (SRMR). Reliability was assessed by administering the form to 20 third- and fourth-year internal medicine trainees. Each item was rated on a Likert scale, and internal consistency was evaluated using Cronbach's alpha.

2.6 Pilot Evaluation

Following confirmation of validity and reliability, the finalized form was applied to evaluate the completeness of admission documentation and medical orders completed by doctors in specialty training. A total of 200 records from 2024 were reviewed, including 100 records completed by first-year trainees in internal medicine, cardiology, and neurology and 100 randomly selected records from trainees in other clinical specialties. The records were obtained from the teaching hospitals of Shahid Faghihi, Namazi, Chamran, Zeynabieh, Dashtgheib, and Khalili. Systematic random sampling was used based on computerised medical records, with one eligible record selected from every 300 patient files at the time of hospital admission. Trainees from multiple specialties, including internal medicine, neurology, cardiology, surgery, obstetrics and gynecology, urology, dermatology, emergency medicine, rehabilitation, orthopedics, psychiatry, pediatrics, and otorhinolaryngology, were evaluated across the five predefined domains. Data were entered into SPSS (version 18) for statistical analysis. Mean scores were calculated for internal medicine, cardiology, and neurology trainees in each domain, as well as an overall mean score for all evaluated trainees.

2.7 Standardization of Scores

Because documentation content varied according to patient condition and diagnosis, raw scores were calculated separately for each trainee in each domain and for the total score. To enable comparison across trainees and specialties, all raw scores were converted to standardized percentage scores out of 100 and used for subsequent analyses.

2.8 Ethical Considerations

The study was approved by the Shiraz University of Medical Sciences Ethics Committee (IR.SUMS.REC.1397.S7168). Each trainee's and patient's privacy was strictly maintained, and no identifiable personal information was recorded or disclosed. All data were securely stored and accessible only to the research team.

3. Results

Results are presented in four sections: (1) development of the multidomain assessment tool, (2) evidence based on content validity obtained through the Delphi process, (3) evidence based on internal structure and internal consistency, and (4) application of the finalized instrument to evaluate admission documentation and medical orders completed by doctors in specialty training.

3.1 Identification of Domains and Items

Based on the literature review and expert panel discussions, an initial evaluation form consisting of five domains was developed to assess the completeness of admission documentation and medical orders. These domains included:

1. Legal components: legibility, signature and stamp, date, time, and ward of admission
2. Initial patient assessment: diagnosis, initial clinical status, activity level, diet type, and allergies
3. Care orders: vital sign monitoring, intravenous fluids, urgent notifications, and nursing instructions
4. Clinical investigations: laboratory, radiologic, and paraclinical studies
5. Medication orders: essential and conditional medications

The literature search yielded 1.345 records, of which 44 were judged relevant to admission documentation and medical order standards (figure 1).

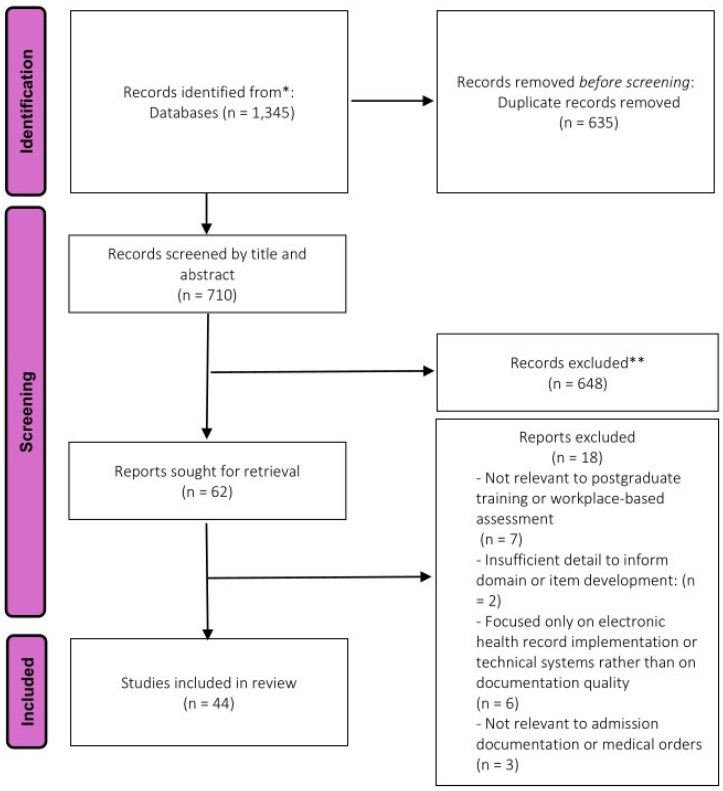

Figure 1. PRISMA-style flow diagram for the initial literature review.

3.2 Delphi Validation of Domains and Items

The Delphi technique was applied to confirm the relevance and content validity of the identified items. In the first Delphi round, 18 of 20 items were accepted, while 2 were revised and reassessed in the second round; after the second round, all items were approved. Expert agreement was high across all domains. All experts (100%) confirmed the necessity and adequacy of items in Domains 3, 4, and 5, while only minimal suggestions for item modification were reported in Domains 1 and 2 (table 1). Agreement among experts was quantified using Kendall's coefficient of concordance (table 2). The coefficient was 0.94 in the first round and 1.00 in the second round, indicating a very high to perfect consensus among the panellists ($p < 0.001$). These findings provide evidence supporting the content validity of the instrument, demonstrating a high level of expert agreement regarding the relevance and completeness of the proposed domains and items.

Table 1. Delphi panel responses on the relevance and adequacy of items across the five domains of the assessment tool.

Domain (Factors Assessed)	Question	Yes, n (%)	No, n (%)
Domain 1 - Legal components (legibility, signature & stamp, date, time, and ward)	Are the items in this domain necessary?	13 (100%)	0
	Are the listed items sufficient?	13 (100%)	0
	Is there a need to add or remove items?	1 (8%)	12 (92%)
Domain 2 - Initial patient assessment (diagnosis, initial status, activity level, diet, and allergies)	Are the items in this domain necessary?	13 (100%)	0
	Are the listed items sufficient?	13 (100%)	0
	Is there a need to add or remove items?	1 (8%)	12 (92%)
Domain 3 - Nursing care orders (vital signs, IV fluids, urgent notifications, and nursing instructions)	Are the items in this domain necessary?	13 (100%)	0
	Are the listed items sufficient?	13 (100%)	0
	Is there a need to add or remove items?	0	13 (100%)
Domain 4 - Clinical investigations (laboratory, radiology, and paraclinical)	Are the items in this domain necessary?	13 (100%)	0
	Are the listed items sufficient?	13 (100%)	0
	Is there a need to add or remove items?	0	13 (100%)
Domain 5 - Medication orders (correct drug name, dose, form, route, and frequency)	Are the items in this domain necessary?	13 (100%)	0
	Are the listed items sufficient?	13 (100%)	0
	Is there a need to add or remove items?	0	13 (100%)

Table 2. Kendall's coefficient of concordance in two Delphi rounds.

Delphi Round	Significance Level (p)	Degrees of Freedom	χ^2	Kendall's W	Number of Experts
First Delphi round	<0.001	14	171.36	0.94	13
Second Delphi round	<0.001	14	182.00	1.00	13

3.3 Validity and reliability of the form

Evidence Based on Internal Structure: based on viewpoints obtained after confirming the evaluation form's content validity using the Delphi method, the completeness of admission documentation and medical orders was examined across five domains: legal components, initial patient assessment, care orders, clinical investigations, and medication orders. These domains comprised 20 items, which were subjected to factor analysis to evaluate the proposed structure. Evidence based on internal structure was examined using CFA. The hypothesised multidomain model showed preliminary support, with RMSEA = 0.063 and CFI = 0.920, suggesting acceptable approximation and comparative model fit, while SRMR = 0.081 indicated acceptable residual-based fit. The chi-square test was statistically significant ($p < 0.001$), a result that should be interpreted cautiously because chi-square is sensitive to sample size and model complexity. Overall, the CFA findings provide initial, but not definitive, support for the proposed multidomain structure. Further testing in larger independent samples, with reporting of additional fit indices and model stability analyses, is needed before the factor structure can be considered robust.

- B1. General Principles in Writing Components Related to Legal Aspects of Medical Orders:
  - B1.1 Having neat and legible handwriting
  - B1.2 Having a signature
  - B1.3 Having a stamp
  - B1.4 Recording the date
  - B1.5 Recording the time
  - B1.6 Specifying the hospitalization ward
- B2. General Principles in Writing Information Related to the Analysis of the Patient's Initial Condition
  - B2.1 Writing the diagnosis
  - B2.2 Writing the initial condition
  - B2.3 Writing the level of activity
  - B2.4 Writing the type of dietary regimen
  - B2.5 Writing allergies
- B3. General Principles of Writing Care Orders
  - B3.1 Checking vital signs
  - B3.2 Administering fluids and medications
  - B3.3 Reporting necessary information
  - B3.4 Nursing care instructions
- B4. General principles for drafting instructions related to clinical evaluations
  - B4.1 Paraclinical investigations
  - B4.2 Radiological examinations
  - B4.3 Laboratory investigations
  - B4.4 Correct documentation of the drug name
  - B4.5 Correct documentation of the drug dosage, pharmaceutical form, and method of administration for medications.

3.4 Internal Consistency

Internal consistency was evaluated using Cronbach's alpha. The overall Cronbach's alpha coefficient was 0.79, indicating acceptable internal consistency for an instrument intended for educational assessment. This finding suggests that the items measure related aspects of documentation performance while retaining sufficient diversity to represent multiple domains.

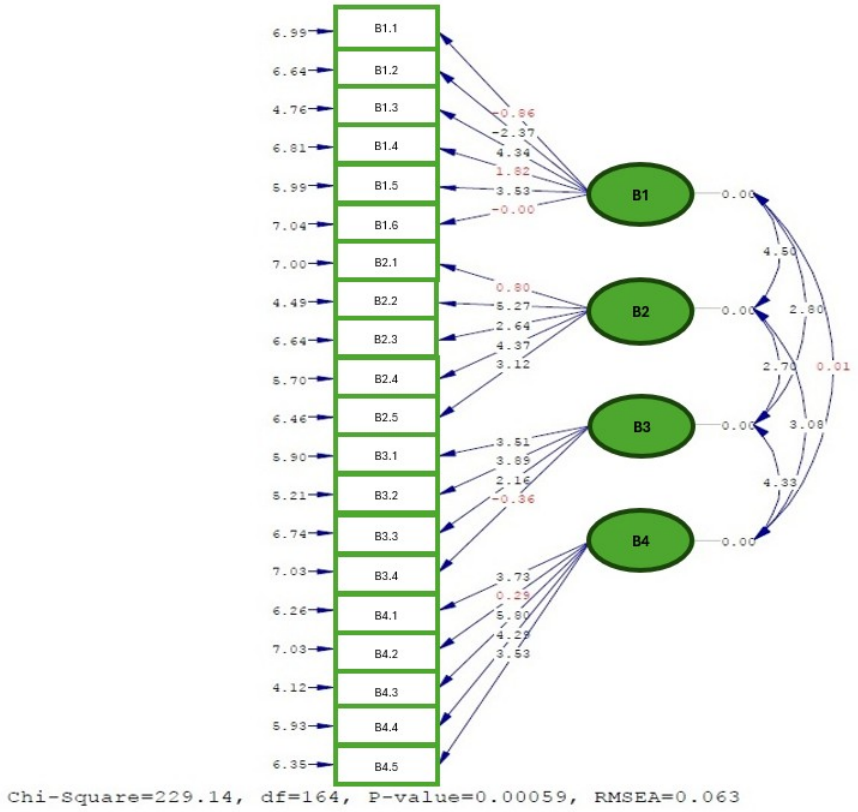

Figure 2. Confirmatory factor analysis of the multidomain evaluation form for admission documentation and medical orders. *Note*. The instrument was developed as a five-domain tool based on a literature review and expert consensus. CFA supported a four-factor empirical model in which clinical investigations and medication orders were combined into a single factor.

3.5 Application of the Instrument

Descriptive statistics were calculated for all trainees across the five domains and for the total score. The overall mean total score was 70.43 ± 7,00, indicating variation in documentation completeness across the evaluated domains. Across domains, the highest mean score was observed for clinical investigations (87.65 ± 19), whereas the lowest was observed for care orders (41.69 ± 15), indicating substantial variation in trainees' documentation performance across domains (table 3).

Table 3. Descriptive statistics of trainees' scores across five domains.

Domain	N	Mean	SD	Minimum	Maximum
Legal components of orders	200	73.12	10	26	95
Initial patient assessment	200	58.27	19	0	80
Nursing care orders	200	41.69	15	5	78
Clinical investigations	168	87.65	19	0	100
Medication orders	137	49.00	19	11	100
Total score	200	70.43	7	37	87

3.6 Domain-Specific Performance by Specialty

Because internal medicine, cardiology, and neurology trainees comprised the three specialty groups included in the psychometric testing phase, domain-specific comparisons were examined in detail for these groups. In the legal components' domain, neurology trainees achieved the highest mean score, followed by internal medicine and cardiology trainees. For initial patient assessment, internal medicine trainees achieved the highest mean score. In contrast, cardiology trainees achieved the highest mean score in the care orders domain. For clinical investigations, neurology trainees showed the highest mean performance, while in the medication orders domain, internal medicine and neurology trainees outperformed cardiology trainees (tables 5-10). Differences were observed among specialties in several documentation domains, although these findings should be interpreted as reflecting variation within the study setting rather than inherent differences between specialties.

Table 5. Mean scores of trainees in the domain of admission documentation and medical orders.

Specialty	N	Mean	SD	Minimum	Maximum
Anaesthesiology	1	68.42	.	68	68
Cardiology	20	62.28	7	52	73
Dermatology	1	68.42	.	68	68
Emergency Medicine	4	57.89	7	52	63
ENT (Ear, Nose, Throat)	13	80.92	9	63	89
Internal Medicine	46	63.69	10	31	84
Neurology	34	69.27	8	47	89
Obstetrics & Gynaecology	18	66.31	5	52	73
Orthopaedics	15	71.57	7	57	89
Paediatrics	17	70.17	10	52	84
Rehabilitation	5	74.73	16	52	89
Psychiatry	3	68.42	7	63	73
Surgery	18	63.15	10	47	73
Urology	5	70.17	8	63	78

Table 6. Mean scores of trainees in the domain of initial patient assessment.

Specialty	N	Mean	SD	Minimum	Maximum
Anaesthesiology	1	0.00	0	0	0
Cardiology	20	61.11	21	20	80
Dermatology	1	60.00	0	60	60
Emergency Medicine	4	60.00	0	60	60
ENT (Ear, Nose, Throat)	13	17.50	7	0	20
Internal Medicine	46	67.92	7	60	80
Neurology	34	68.10	17	0	80
Obstetrics & Gynaecology	18	60.00	0	60	60
Orthopaedics	15	26.66	9	20	40
Paediatrics	17	60.00	0	60	60
Rehabilitation	5	64.00	8	60	80
Psychiatry	3	80.00	0	80	80
Surgery	18	57.50	19	20	80
Urology	5	60.00	0	60	60

Table 7. Mean scores of trainees in the domain of nursing care orders.

Specialty	N	Mean	SD	Minimum	Maximum
Anaesthesiology	1	0.00	0	0	0
Cardiology	19	58.28	13	33	81
Dermatology	1	50.00	0	50	50
Emergency Medicine	4	56.94	1	55	58
ENT (Ear, Nose, Throat)	13	36.79	12	20	57
Internal Medicine	45	57.43	10	25	72
Neurology	30	49.35	9	33	66
Obstetrics & Gynaecology	18	56.89	14	25	72
Orthopaedics	15	52.65	7	40	70
Paediatrics	17	54.35	7	44	70
Rehabilitation	5	24.76	24	0	57
Psychiatry	10	46.48	9	33	60
Surgery	5	57.14	10	50	64
Urology	5	49.36	12	40	63

Table 8. Mean Scores of trainees in the Clinical Assessment Domain.

Specialty	N	Mean	SD	Minimum	Maximum
Cardiology	14	84.25	25	11	100
Dermatology	1	100	0	100	100
Emergency Medicine	4	83.33	0	83	83
Otorhinolaryngology (ENT)	6	100	0	100	100
Internal Medicine	46	84.61	10	66	100
Neurology	34	90.84	13	44	100
Obstetrics & Gynaecology	17	77.77	16	66	100
Orthopaedics	15	98.14	4	83	100
Paediatrics	17	81.94	16	66	100
Rehabilitation Medicine	5	90	22	50	100
Psychiatry	5	91.66	11	83	100
Surgery	18	90.27	23	33	100
Urology	5	94.44	9	83	100

Table 9. Mean Scores of trainees in the Medication Orders Domain.

Specialty	N	Mean	SD	Minimum	Maximum
Anaesthesiology	1	66.66	0	66	66
Cardiology	16	68.98	16	27	83
Dermatology	1	91.66	0	91	91
Otorhinolaryngology (ENT)	12	84.25	10	66	100
Internal Medicine	45	77.41	12	33	94
Neurology	28	77.38	15	33	100
Obstetrics & Gynaecology	5	91.66	8	83	100
Orthopaedics	15	80.09	13	50	100
Paediatrics	15	66.28	10	50	83
Rehabilitation Medicine	5	91.57	6	83	100
Psychiatry	2	43.33	0	43	43
Surgery	7	70	24	40	100
Urology	2	83.33	0	83	83

Table 10. Mean total score of trainees in completing admission documentation and medical orders.

Specialty	N	Mean	SD	Minimum	Maximum
Anaesthesiology	1	56,66	0	56	56
Cardiology	20	65,60	8	49	77
Dermatology	1	72,34	0	72	72
Emergency Medicine	4	61,63	3	58	64
Otorhinolaryngology (ENT)	13	68,48	6	60	76
Internal Medicine	46	73,43	4	61	81
Neurology	34	72,90	7	48	84
Obstetrics & Gynaecology	18	64,78	6	53	75
Orthopaedics	15	70,24	5	56	78
Paediatrics	17	66,03	4	57	71
Rehabilitation Medicine	5	77,76	16	47	86
Psychiatry	3	67,36	8	61	73
Surgery	18	67,09	12	49	92
Urology	5	69,33	8	60	74

3.6 Comparison of scores across internal medicine, cardiology, and neurology

Descriptive statistics for internal medicine, cardiology, and neurology trainees were calculated separately for each domain and for the total score. As shown in tables 11-13, neurology trainees obtained higher mean scores in Domains 1, 4, and 5; cardiology trainees scored higher in Domain 3; and internal medicine trainees scored higher in Domain 2 and in the total score. Analysis of variance showed statistically significant differences among the three specialty groups in Domains 1, 2, and 3, as well as in the total score ($p \leq 0.05$). No statistically significant differences were observed in Domains 4 or 5 ($p \geq 0.05$).

Table 11. Mean Scores of Internal Medicine, Cardiology, and Neurology Residents Across Five Domains and Total Score.

Specialty	Statistic	Legal Aspects of Orders	Initial Situation Analysis	Nursing Orders	Clinical Assessments	Medication Orders	Total Score
Cardiology	N	20	20	19	14	16	20
	Mean ± SD	62 ± 7	61 ± 21	58 ± 13	84 ± 25	68 ± 16	65 ± 8
Internal Medicine	N	46	46	45	46	45	46
	Mean ± SD	63 ± 10	76 ± 7	57 ± 10	84 ± 10	76 ± 12	73 ± 4
Neurology	N	34	34	30	34	28	34
	Mean ± SD	69 ± 8	68 ± 17	49 ± 9	90 ± 13	77 ± 15	72 ± 7

Table 12. Results of ANOVA Comparing Mean Scores of Internal Medicine, Cardiology, and Neurology Trainees.

Scores	F Statistic	Significance Level (p-value)
Legal Aspects of Orders	5.73	0.005
Initial Situation Analysis	5.48	0.006
Nursing Care Orders	4.79	0.011
Clinical Assessments	1.66	0.197
Medication Orders	0.31	0.731
Total Score	9.50	< 0.001

Statistically significant differences were identified in Domains 1–3 and in the total score, whereas no significant differences were observed for the Clinical Investigations or Medication Orders domains.

Table 13. Results of the Scheffé Post Hoc Test for Determining Differences in Mean Scores among Internal Medicine, Cardiology, and Neurology Trainees (* statistically significant at $p \leq 0.05$).

Dependent Variable	Specialty (I)	Specialty (J)	I–J	95% CI	P level
Legal Aspects of Orders	Cardiology	Internal Medicine	-2	-9 to 4	0.71
		Neurology	**-8***	-15 to -1	0.01
	Internal Medicine	Cardiology	2	-4 to 9	0.71
		Neurology	-5	-11 to 0.7	0.05
	Neurology	Cardiology	**8***	1 to 15	0.01
		Internal Medicine	5	-7 to 11	0.05
Initial Situation Analysis	Cardiology	Internal Medicine	**-15***	-27 to -3	<0.001
		Neurology	-6	-18 to 4	0.31
	Internal Medicine	Cardiology	**15***	3 to 27	<0.001
		Neurology	8	-1 to 18	0.10
	Neurology	Cardiology	6	-4 to 18	0.35
		Internal Medicine	-8	-18 to 1	0.10
Nursing Care Orders	Cardiology	Internal Medicine	-2	-12 to 7	0.81
		Neurology	7	-2 to 16	0.15
	Internal Medicine	Cardiology	2	-7 to 12	0.81
		Neurology	**9***	1 to 18	0.01
	Neurology	Cardiology	-7	-16 to 2	0.15
		Internal Medicine	**-9***	-18 to -1	0.01
Clinical Assessments	Cardiology	Internal Medicine	-0.3	-13 to 12	0.99
		Neurology	-6	-19 to 5	0.42
	Internal Medicine	Cardiology	0.35	-12 to 13	0.99
		Neurology	-6	-15 to 3	0.27
	Neurology	Cardiology	6	-5 to 19	0.42
		Internal Medicine	6	-3 to 15	0.27
Medication Orders	Cardiology	Internal Medicine	-2	-16 to 12	0.93
		Neurology	-4	-18 to 9	0.74
	Internal Medicine	Cardiology	2	-12 to 16	0.93
		Neurology	-2	-14 to 9	0.89
	Neurology	Cardiology	4	-9 to 18	0.74
		Internal Medicine	2	-9 to 14	0.89
Total Score	Cardiology	Internal Medicine	**-8***	-13 to -3	<0.001
		Neurology	**-7***	-12 to -2	<0.001
	Internal Medicine	Cardiology	**8***	3 to 13	<0.001
		Neurology	0.7	-3 to 4	0.91
	Neurology	Cardiology	**7***	2 to 12	<0.001
		Internal Medicine	-0.7	-4 to 3	0.91

Scheffé post hoc testing indicated that the difference in Domain 1 was between cardiology and neurology trainees; in Domain 2, between cardiology and internal medicine trainees; and in Domain 3, between internal medicine and neurology trainees. For the total score, significant differences were

observed between cardiology and internal medicine trainees and between cardiology and neurology trainees ($p \leq 0.05$).

3.7 Determining the Correlation Between Domain Scores and the Total Score

Normality of scores across all five domains was confirmed; therefore, Pearson's correlation coefficient was used to examine associations between domain scores and the total score. According to table 14, the correlation coefficients between all domains and the total score were positive and statistically significant ($p < 0.001$), providing additional evidence that each domain contributed meaningfully to the overall assessment score. The correlations among the domains themselves were also examined. Significant correlations were observed between Domains 1 and 2, Domains 3 and 4, and Domains 4 and 5 (table 15).

Table 14. Results of Normality Testing of Residents' Scores Using the One-Sample Kolmogorov–Smirnov Test.

Domain	Number	Mean Score	SD	Kolmogorov–Smirnov Z	Significance Level (p)
Legal Aspects of Orders	200	73.12	10	2.29	<0.001
Initial Situation Analysis	200	58.27	19	4.64	<0.001
Nursing Care Orders	200	41.69	15	3.00	<0.001
Clinical Assessments	168	87.65	19	4.86	<0.001
Medication Orders	137	49.00	19	2.63	<0.001
Total Score	200	70.43	7	0.86	0.43

* Based on the test results, the distribution of scores was normal.

Table 15. Results of Correlation Analysis Between the Five Domains and the Total Score ([a], $p < 0.001$; [b], $p \leq 0.05$).

	Total Score	Legal Aspects of Orders	Initial Situation Analysis	Nursing Care Orders	Clinical Assessments	Medication Orders
Total Score	1	0.49[a]	0.31[a]	0.20[a]	0.46[a]	0.57[a]
p-value		<0.001	<0.001	<0.001	<0.001	<0.001
Legal Aspects of Orders	0.49[a]	1	-0.02	-0.18[a]	0.13[b]	-0.00
p-value	<0.001		0.67	<0.001	0.03	0.90
Initial Situation Analysis	0.31[a]	-0.02	1	0.20[a]	-0.14[b]	0.03
p-value	<0.001	0.67		<0.001	0.01	0.61
Nursing Care Orders	0.20[a]	-0.18[a]	0.20[a]	1	0.05	0.04
p-value	<0.001	<0.001	<0.001		0.40	0.55
Clinical Assessments	0.46[a]	0.13[b]	-0.14[b]	0.05	1	0.17[a]
p-value	<0.001	0.03	0.01	0.40		<0.001
Medication Orders	0.57[a]	-0.00	0.03	0.04	0.17[a]	1
p-value	<0.001	0.90	0.61	0.55	<0.001	

4. Discussion

This study developed a structured, multidomain tool to assess the completeness of admission documentation and medical orders completed by doctors in specialty training and gathered initial validity evidence for its use in postgraduate clinical education. Admission documentation is a routine workplace activity, but it also reflects important aspects of clinical practice, including information synthesis, communication, professional accountability, and attention to patient safety. In this sense, documentation can provide supervisors with an accessible product of clinical work that may be used to support feedback and identify learning needs (15). Consistent with recent work in medical education assessment tool development, the present study highlights the importance of combining expert input with empirical validation before assessment instruments are used beyond formative purposes (16).

The need for such assessment tools is supported by a growing body of literature demonstrating persistent challenges in the quality and completeness of clinical documentation. Previous studies have consistently shown ongoing challenges in the completeness and quality of medical documentation across clinical settings. Incomplete summaries, inconsistent documentation practices, and non-compliance with established standards have all been reported, with potential implications for continuity and safety of care (17-19). Other studies suggest that structured documentation formats and electronic medical record systems can improve completeness, although the magnitude of improvement varies by context and implementation (15, 20). These findings support the need for structured approaches to documentation, assessment, and feedback.

The tool includes five domains: legal components, initial patient assessment, care orders, clinical investigations, and medication orders. These domains were derived from the literature review and expert consensus and were intended to capture practical components of documentation completeness in the local admission documentation context. The Delphi findings provide content-related evidence that the domains and items were considered relevant and appropriate by experienced faculty members. However, this should be interpreted as one source of validity evidence rather than proof that the instrument is fully validated for all assessment purposes.

The CFA findings provided preliminary support for the proposed multidomain structure, with RMSEA = 0.063, CFI = 0.920, and SRMR = 0.081 suggesting acceptable model fit. Nevertheless, the internal structure evidence should be interpreted cautiously. The CFA sample size was modest for a 20-item multidomain instrument. Future studies with larger and independent samples are needed to examine model stability and provide more complete psychometric evidence.

A further issue relates to the alignment between the conceptually defined domains and their representation in the CFA results. Although the tool was developed as a five-domain instrument, with clinical investigations and medication orders specified as separate domains, the available CFA output grouped these items within a single factor. This finding suggests that the distinction between these domains may not have been fully reflected in the current sample's empirical structure. Therefore, the CFA findings should be interpreted as preliminary evidence of internal structure. Future studies should evaluate the proposed five-domain model in larger and more diverse samples and report additional evidence regarding model stability and factor structure.

When applied to trainee documentation, the tool identified variation in documentation completeness across domains and specialties. These differences may reflect variation in clinical routines, supervisory expectations, documentation cultures, or exposure to specific admission tasks. Because the study was not designed to determine the causes of specialty differences, these findings should not be interpreted as evidence of superior or inferior specialty performance. Instead, they

suggest areas where documentation expectations may need clarification and where targeted teaching or feedback may be useful.

At this stage, the most appropriate use of the tool is for formative purposes. It may help supervisors provide more specific feedback by identifying which components of admission documentation and medical orders are complete and which require improvement. For example, rather than giving broad feedback such as "improve documentation", supervisors could use the domain scores to discuss omissions in legal components, patient assessment, care orders, investigations, or medication orders. Claims about use for progression decisions, summative assessment, or longitudinal competence tracking require further evidence, particularly regarding inter-rater reliability, response processes, relationships with other assessment measures, and consequences of use.

The study has several limitations. First, it was conducted in teaching hospitals within a single national and institutional context, and the Iranian admission documentation format may differ from those used in other countries. This may limit transferability, although the domains are broadly relevant to clinical documentation and order writing. Second, the tool assesses completeness of documentation rather than directly measuring the quality of clinical reasoning, appropriateness of management decisions, patient outcomes, or feedback uptake. Third, the study did not examine inter-rater reliability, inter-rater agreement, generalizability, or temporal stability, which are important for any instrument intended for workplace-based assessment. Fourth, evidence related to response processes, relationships with other variables, and consequences of use was not collected. These issues should be addressed before the tool is used for high-stakes or summative decisions.

Despite these limitations, the study offers a structured starting point for assessing an under-evaluated workplace activity in specialty training. This tool may be useful for formative feedback, documentation audits, and identifying areas for educational improvement. Further research should test the instrument in larger, more diverse samples, compare it with existing documentation assessment tools, examine assessor training requirements, and evaluate whether its use improves feedback quality, trainee documentation practices, or patient care processes over time.

5. Conclusions

- A multidomain tool was developed to assess admission documentation and medical orders in specialty training.
- Initial findings support its content validity, internal structure, and internal consistency.
- The tool may help supervisors identify documentation gaps and provide structured formative feedback.
- Further validation is needed before use in high-stakes or summative assessment.

Funding: This research received no external funding.

Declaration of conflict of interest: The authors declare no conflicts of interest.

Author contributions: MH: conceptualisation, methodology, formal analysis, validation, supervision, drafting of the original manuscript, and manuscript review. MA: research, data acquisition, and manuscript review. RR: research, data acquisition, and manuscript review. MH: research, data acquisition, and manuscript review. SJ: research, formal analysis, supervision, critical review of the manuscript, drafting, and manuscript review. All authors contributed to the interpretation of the results, reviewed and approved the final version of the manuscript, and accepted responsibility for all aspects of the work.

6. References.

1. Carraccio C, Englander R, Van Melle E, et al. Advancing competency-based medical education: a charter for clinician–educators. *Acad Med*. **2016**, 91(5), 645-649. https://doi.org/10.1097/ACM.0000000000001048
2. Schuwirth LW, van der Vleuten CP. A history of assessment in medical education. *Adv Health Sci Educ Theory Pract*. **2020**, 25(5), 1045-1056. https://doi.org/10.1007/s10459-020-10003-0
3. Weller JM, Coomber T, Chen Y, et al. Key dimensions of innovations in workplace-based assessment for postgraduate medical education: a scoping review. *Br J Anaesth*. **2021**, 127(5), 689-703. https://doi.org/10.1016/j.bja.2021.06.038
4. Demsash AW, Kassie SY, Dubale AT, et al. Health professionals' routine practice documentation and its associated factors in a resource-limited setting: a cross-sectional study. *BMJ Health Care Inform*. **2023**, 30(1), e100699. https://doi.org/10.1136/bmjhci-2022-100699
5. Wilesmith S, Mandrusiak A, Lang R, et al. Educational interventions to develop and enhance clinical documentation skills in health professional students: a systematic review. *Clin Teach*. **2025**, 22(5), e70157. https://doi.org/10.1111/tct.70157
6. Grauer A, Rosen A, Applebaum JR, et al. Examining medication ordering errors using AHRQ network of patient safety databases. *J Am Med Inform Assoc*. **2023**, 30(5), 838-845. https://doi.org/10.1093/jamia/ocad007
7. Richardson KM, Cristiano JA, Schafer KR, et al. Writing is thinking: implementation and evaluation of an internal medicine residency clinical reasoning and documentation curriculum. *Med Sci Educ*. **2022**, 32(4), 773-777. https://doi.org/10.1007/s40670-022-01570-5
8. Bowker D, Torti J, Goldszmidt M. Documentation as composing: how medical students and residents use writing to think and learn. *Adv Health Sci Educ Theory Pract*. **2023**, 28(2), 453-475. https://doi.org/10.1007/s10459-022-10167-x
9. Burke HB, Hoang A, Becher D, et al. QNOTE: an instrument for measuring the quality of EHR clinical notes. *J Am Med Inform Assoc*. **2014**, 21(5), 910-916. https://doi.org/10.1136/amiajnl-2013-002321
10. Stetson PD, Bakken S, Wrenn JO, et al. Assessing electronic note quality using the physician documentation quality instrument (PDQI-9). *Appl Clin Inform*. **2012**, 3(2), 164-174. https://doi.org/10.4338/ACI-2011-11-RA-0070
11. Baker EA, Ledford CH, Fogg L, et al. The IDEA assessment tool: assessing the reporting, diagnostic reasoning, and decision-making skills demonstrated in medical students' hospital admission notes. *Teach Learn Med*. **2015**, 27(2), 163-173. https://doi.org/10.1080/10401334.2015.1011654
12. Kotwal S, Klimpl D, Tackett S, et al. Documentation of clinical reasoning in admission notes of hospitalists: validation of the CRANAPL assessment rubric. *J Hosp Med*. **2019**, 14(12), 746-753. https://doi.org/10.12788/jhm.3233
13. Kelly WF, Hawks MK, Johnson WR, et al. Assessment tools for patient notes in medical education: a scoping review. *Acad Med*. **2025**, 100(3), 358-374. https://doi.org/10.1097/ACM.0000000000005886
14. Messick S. Validity. In: Linn RL, ed. Educational Measurement. 3rd ed. *American Council on Education and Macmillan Publishing Company*. **1989**, 13-103.
15. Ebbers T, Kool RB, Smeele LE, et al. The impact of structured and standardized documentation on documentation quality: a multicenter, retrospective study. *Journal of Medical Systems.* **2022**, 46(7), 46. https://doi.org/10.1007/s10916-022-01837-9
16. Sánchez Ferrer F, Díaz J, Maestre Ferrándiz J, Cortes-Castell E, Gil Guillen V, Rizo M, Nso-Roca AP. Prescribing Patient Associations: Bringing Medical Education Closer to the Patient's Reality. *Spanish Journal of Medical Education*. **2026**, 7(5). https://doi:10.6018/edumed.717851
17. Dean B, Barber N, Schachter M. What is a prescribing error? *BMJ Quality & Safety*. **2000**, 9(4), 232-237. https://doi.org/10.1136/qhc.9.4.232

18. Usama M, Arshad R, Rizwan U, et al. Evaluating the Quality and Completeness of Inpatient Medical Record Documentation: A Clinical Audit. *Journal of Medical & Health Sciences Review*. **2025**, 2(1). https://doi.org/10.62019/4m4e3720
19. Tariq MD, Hamza HM, Ahsan A, et al. Unveiling documentation deficiencies: a clinical audit of histopathology request forms at a tertiary care hospital. *Ann Med Surg (Lond)*. **2025**, 87(4), 2000-2005. https://doi.org/10.1097/MS9.0000000000003038
20. Wurster F, Herrmann C, Beckmann M, et al. Differences in changes of data completeness after the implementation of an electronic medical record in three surgical departments of a German hospital: a longitudinal comparative document analysis. BMC Medical Informatics and Decision Making. **2024**, 24(1), 258. https://doi.org/10.1186/s12911-024-02667-0