

FIELD TESTING AND PRELIMINARY ITEM ANALYSIS OF THE PATIENT ASSESSMENT AND MANAGEMENT TOOL–PALLIATIVE CARE (PAMT-PC): A 37-ITEM KNOWLEDGE QUESTIONNAIRE AMONG HEALTHCARE PROFESSIONALS IN ANDHRA PRADESH, INDIA

Shahedha Parveen^{1*}, S. Vinay Kumar², B. Sneha Latha², P. Sophia³, Athaluri Vishnu Vardhan⁴, John Dogulas Palleti⁵

^{1*}Professor, Rangaraya Medical College, Kakinada, Andhra Pradesh, India. drshahedha@gmail.com

²Siddhartha Medical College, Vijayawada, Andhra Pradesh, India (S. Vinay Kumar: Professor and Head of the Department; B. Sneha Latha: Associate Professor). vinay67@yahoo.com; snehabhashyam@gmail.com

³Professor and Head of the Department, Government Medical College, Rajamahendravaram, Andhra Pradesh, India. dr.sophia29@gmail.com

⁴Director of Medical Education, Andhra Pradesh, Vijayawada, India. athalurivishnu@gmail.com

⁵Chief Scientist, Centre for Computational and Biological Sciences (CCBS), Visakhapatnam, Andhra Pradesh, India. drjohndpalleti@gmail.com

*Corresponding author: Dr Shahedha Parveen, India. Email: drshahedha@gmail.com

ABSTRACT

Background: Healthcare professionals in India lack validated, context-specific instruments for assessing palliative care knowledge. The only available Indian tool — the Palliative Care Knowledge Questionnaire-Basic (PCKQ-B)— was developed exclusively for Delhi-NCR nursing staff, lacks factor analysis, and omits Communication Skills and Procedural Skills domains mandated by India's Competency-Based Medical Education (CBME 2019) curriculum. This study reports the development, content validation, pilot testing, field testing, and test-retest reliability assessment of the Patient Assessment and Management Tool- Palliative Care (PAMT-PC), a 37-item, seven-domain palliative care knowledge questionnaire for multidisciplinary healthcare professionals in Andhra Pradesh, India.

Methods: A four-phase, COSMIN-guided instrument development study was conducted: (1) systematic item generation from five existing tools and CBME 2019 competencies; (2) expert panel content validity assessment (n = 9 experts, two Delphi rounds); (3) pilot testing (n = 30) with cognitive debriefing; and (4) field testing on n = 230 healthcare professionals (173 nurses, 33 doctors, 24 allied health workers) at Government General Hospital (GGH) Kakinada, Andhra Pradesh; all four phases were completed sequentially between December 2023 and March 2025. Test-retest reliability was assessed in a sub-sample of n = 45 using supervised paper-based administration at a two-week interval.

Results: Content validity: S-CVI/Ave = 0.868; mean I-CVI = 0.868 (range 0.778–1.000); mean κ^* = 0.860; 15 items revised across two Delphi rounds; all 37 items achieved I-CVI ≥ 0.778 . Pilot testing: median completion time 19.8 min; preliminary KR-20 = 0.551; five items revised. Field test (n = 230): mean score $23.78 \pm 4.54/37$ (64.3%); KR-20 = 0.720 (acceptable); mean inter-item correlation = 0.063; 22/37 items retained; 12 flagged for review; three items (Q23, Q24, Q32a) showed negative point-biserial correlations and require revision. EFA sampling adequacy: KMO = 0.64; Bartlett's $\chi^2(666) = 1,768.4$, $p < 0.001$. Test-retest (n = 45): ICC(2,1) = 0.907 (95% CI: 0.837–0.947; excellent); SEM = 1.24 score points; Bland–Altman limits of agreement –3.70 to +3.53. Known-groups validity: prior training significantly predicted higher scores ($t(195) = 3.47$, $p < 0.001$, Cohen's $d = 0.626$).

Conclusion: The PAMT-PC demonstrated near-acceptable content validity, acceptable internal consistency (KR-20 = 0.720), and excellent test-retest stability (ICC = 0.907) in this preliminary validation sample. Three items require immediate revision and 12 require review. Multi-district validation (n ≥ 500) and confirmatory factor analysis are the recommended next steps toward full psychometric validation.

KEYWORDS: Questionnaire development; palliative care knowledge; item analysis; KR-20; content validity; test-retest reliability; health professionals; India; Andhra Pradesh; COSMIN; PAMT-PC

1. INTRODUCTION

Palliative care is defined by the World Health Organization (WHO) as an approach that improves the quality of life of patients and families facing life-threatening illness through prevention and relief of suffering [1]. Globally, over 56.8 million people require palliative care annually, with 78% residing in low- and middle-income countries [2]. The Lancet Commission on Palliative Care and Pain Relief estimated that 61 million people experience serious health-related suffering annually, with over 80% in low-resource settings [3]. India faces an estimated 7–10 million people requiring palliative care annually, yet fewer than 4% receive it [4]. India ranked 67th of 80 countries on the Economist Intelligence Unit Quality of Death Index 2015, with particularly weak scores in affordability and human resource capacity [5].

The National Programme for Palliative Care (NPPC), launched in 2012, mandated at least four palliative care beds per district hospital [6]. The 2019 Competency-Based Medical Education (CBME) reform incorporated palliative care competencies across anaesthesia, pharmacology, surgery, and community medicine curricula; however, documented

deficits in dedicated faculty, assessment, and structured training persist [7]. Available reports suggest that palliative care services in Andhra Pradesh remain limited and unevenly distributed across districts.

Suboptimal palliative care knowledge among healthcare professionals is a principal driver of India's end-of-life care crisis. Prem et al. reported a mean Palliative Care Knowledge Test (PCKT) score of 35.8% among 363 South Indian nurses [8]. Sadhu et al. found that 89% of medical undergraduates believed morphine causes addiction in terminally ill patients [9]. Gupta et al. documented a pre-training Palliative Care Quiz for Nursing (PCQN) mean of 42.3% in an ELNEC-India study [10]. No published systematic study of palliative care knowledge gaps among healthcare professionals specifically in Andhra Pradesh existed prior to the present investigation.

The two most widely used international palliative care knowledge instruments are the PCQN [11] (20 items; KR-20 = 0.78) and PCKT [12] (20 items; KR-20 = 0.81), both developed in high-income settings without Indian contextual validation. The only Indian-validated tool, the PCKQ-B [13] (25 items; KR-20 = 0.65), was developed exclusively for Delhi-NCR nursing staff and includes neither Communication Skills nor Procedural Skills — competencies required by the NMC CBME 2019 framework. Anaesthesiologists play a central but under-recognised role in palliative care: managing cancer pain, providing palliative sedation, making end-of-life ventilation decisions, and supporting families on withdrawal of treatment [7]. The PAMT-PC was developed to fill this gap.

The aim was to develop and validate the PAMT-PC, a 37-item, seven-domain palliative care knowledge questionnaire for multidisciplinary healthcare professionals in Andhra Pradesh. Objectives were: (1) generate and content-validate items through expert panel review; (2) pilot-test the instrument and assess feasibility; (3) administer the instrument to a field-test sample ($n = 230$) and assess item-level psychometric properties; (4) evaluate internal consistency using KR-20; (5) assess test-retest reliability using ICC; (6) assess EFA sampling adequacy; and (7) test known-groups construct validity.

2. MATERIALS AND METHODS

2.1 Study Design and Reporting Standards: This four-phase, mixed-methods instrument development study is reported against the COSMIN Reporting Guideline v2.0 [14], the GRRAS guideline [15] for reliability assessment, and the STROBE guideline [16] for the cross-sectional field-testing component. COSMIN domains — content validity, structural validity, internal consistency, reliability, and construct validity — were applied to the PAMT-PC as an edumetric knowledge test, consistent with Nakazawa et al. [12]. The AERA/APA/NCME Standards for Educational and Psychological Testing [17] provided complementary guidance for item-level requirements.

2.2 Study Setting and Population: Field testing was conducted at Government General Hospital (GGH) Kakinada, affiliated with Rangaraya Medical College, East Godavari District, Andhra Pradesh — a 900-bed tertiary care teaching hospital under Dr. NTR University of Health Sciences. The four study phases were conducted sequentially between December 2023 and March 2025. Phases 1 and 2 (item generation with a test blueprint, followed by the two-round expert content validity review) were completed between December 2023 and the start of the pilot in November 2024; Phase 3 (supervised paper-based pilot testing with cognitive debriefing) was conducted in November–December 2024; and Phase 4 (online field testing) began after the pilot revisions were finalised and was completed in March 2025, during which the two supervised paper-based test-retest administrations (14 days apart) were also completed. Each phase began only after the preceding phase was complete, so all field-test participants received the final, pilot-revised 37-item instrument. Eligible participants were healthcare professionals directly involved in patient care: medical doctors (MBBS and above), nurses (GNM/BSc Nursing and above, including student nurses with ≥ 6 month's clinical exposure), and allied health workers (physiotherapists and anaesthesia technicians). Palliative care specialists and IAPC-certified palliative nurses were excluded to avoid ceiling effects [13].

2.3 Instrument Description and Scoring: The PAMT-PC contains 35 numbered questions (Q1–Q35). Questions Q13 and Q32 each contain two independently scored sub-parts (a and b), yielding 37 scoreable items and a maximum total score of 37. Items span seven clinical domains: Philosophy (1 item), Pain (14 items), GIT/Nutrition (5 items), Psychiatry (6 items), Dyspnoea (3 items), Procedure Skills (3 items), and Communication Skills (5 items). Each item uses a three-option response format (Yes / No / Don't Know). Scoring: correct response = 1; incorrect response or Don't Know = 0. Total score range = 0–37. This scoring convention is consistent with the PCKT [12] and PCKQ-B [13].

2.4 Phase 1: Item Generation: Items were generated deductively from five existing palliative care knowledge tools (PCQN, PCKT, PEACE-Q, PCKQ-B, PaCKS) and inductively from WHO palliative care guidelines, IAPC clinical protocols, and NMC CBME 2019 competency frameworks. A test blueprint mapped 40 candidate items to seven clinical domains and three Bloom's revised taxonomy cognitive levels — Remember, Understand, and Apply [18]. Two new domains absent from all existing tools were developed: Procedural Skills and Communication Skills, addressing CBME 2019 competency gaps [7]. Following DeVellis and Thorpe [19], the initial pool was approximately three times the desired final length. Questions Q13 and Q32 were split into sub-parts to capture two clinically distinct knowledge components within each question.

2.5 Phase 2: Content Validity (Expert Panel): A nine-member expert panel reviewed all 37 items across two Delphi rounds. Panel size followed Lynn [21] and Polit and Beck [20]. The panel comprised two IAPC-certified palliative care specialists, two senior anaesthesiologists (>10 years), one palliative care nurse educator, one medical educator (FAIMER fellowship), one oncologist, one biostatistician, and one clinical pharmacologist. Each expert rated item relevance on a four-point Likert scale (1 = not relevant to 4 = highly relevant; Lynn [21]). Item Content Validity Index (I-CVI) was calculated as the proportion of experts rating ≥ 3 ; Scale Content Validity Index/Average (S-CVI/Ave) as the mean of all I-CVIs; and modified kappa (κ^*) corrected I-CVI for chance agreement (Polit et al. [22]). Acceptable thresholds: I-CVI ≥ 0.78 with ≥ 9 experts; S-CVI/Ave ≥ 0.90 ; $\kappa^* \geq 0.74$ = excellent [22]. Items with I-CVI < 0.78 in Round 1 were revised and re-evaluated in Round 2. None of the experts was an author or supervisor of this study.

2.6 Phase 3: Pilot Testing: Pilot testing was conducted with $n = 30$ healthcare professionals (10 doctors, 14 nurses, 6 allied health workers; mean age 29.6 ± 4.5 years) at an associated hospital of GGH Kakinada that was not included in the field study, using supervised paper-based administration. Cognitive debriefing interviews (Willis [23]) were conducted individually with a sub-sample of $n = 10$ pilot participants (all doctors) to assess item comprehension and cultural appropriateness. Outcomes assessed: median completion time, item-level difficulty, and preliminary KR-20. Five items required wording revision following pilot feedback (Section 3.3).

2.7 Phase 4: Field Testing: Field testing was conducted on 230 eligible healthcare professionals recruited by voluntary convenience sampling at GGH Kakinada. Of 231 Google Forms submissions received, 230 were eligible after excluding one House surgeon who did not meet the minimum clinical exposure criterion. The instrument was self-administered online via Google Forms without supervision, and data collection was anonymous. Participants were instructed to respond based on current knowledge without consulting references.

2.8 Item Analysis: For each of the 37 items: (a) Difficulty index (p) = proportion answering correctly (optimal: 0.30–0.70; acceptable: 0.20–0.90 [24]); (b) Discrimination index (D) = $p_{upper27\%} - p_{lower27\%}$ (≥ 0.20 acceptable; ≥ 0.40 excellent [24]); (c) Corrected point-biserial correlation (r_{pb}) between item and item-corrected total (≥ 0.20 acceptable; ≥ 0.30 preferred [25]); (d) KR-20 if item deleted. Items with negative r_{pb} were flagged as Revise; items with $D < 0.20$ or $r_{pb} < 0.20$ were flagged as Review; items meeting both criteria were retained.

2.9 Internal Consistency — KR-20: KR-20 [26] was calculated overall and per domain. For dichotomous items, KR-20 numerically equals Cronbach's α [27]; KR-20 is reported as it is the convention in the palliative care knowledge tool literature. Acceptable threshold: KR-20 ≥ 0.70 ; good: ≥ 0.80 [25]. Mean inter-item correlation (IIC) was calculated; acceptable range for multidimensional knowledge tests: 0.10–0.40 [28]. Low mean IIC with acceptable overall KR-20 is expected and appropriate for broad-domain knowledge tests [29].

2.10 Test-Retest Reliability: Test-retest reliability was assessed in $n = 45$ healthcare professionals (34 nurses, 7 doctors, 4 allied health workers; mean age 27.8 ± 4.7 years) recruited by convenience from the field-test cohort at GGH Kakinada. These participants completed the PAMT-PC on two further occasions 14 days apart (T1 and T2), using supervised paper-based administration with identical instructions; T1 and T2 responses were linked through anonymous participant codes. No palliative care training sessions occurred during the interval. ICC(2,1) — two-way random effects, absolute agreement, single measures [30] — was calculated with 95% CI per the GRRAS guideline [15]. Standard Error of Measurement ($SEM = SD \times \sqrt{1 - ICC}$) and Bland–Altman analysis were performed. ICC values were interpreted as poor (< 0.50), moderate (0.50–0.74), good (0.75–0.89), and excellent (≥ 0.90), following Koo and Li [30]. Agreement was considered acceptable if both Bland–Altman limits of agreement fell within ± 5 score points (13.5% of the scale range), a benchmark defined before analysis. Test-retest reliability was estimated for the total score only.

2.11 Structural Validity — EFA Sampling Adequacy: EFA sampling adequacy was evaluated using the Kaiser-Meyer-Olkin (KMO) measure (acceptable ≥ 0.60 [31]) and Bartlett's test of sphericity (significant $p < 0.001$). Detailed EFA factor loading interpretation was deferred because the field-test sample was modest for stable multidimensional factor modelling. Full EFA on a tetrachoric correlation matrix using WLS extraction and Promax rotation (R package psych [33]), and subsequent CFA using WLSMV estimation with model fit judged against established cut-offs [32], are planned for a future larger validation study ($n \geq 500$).

2.12 Known-Groups Construct Validity: Three pre-specified known-groups hypotheses were tested per COSMIN methodology [34]: H1: Healthcare professionals with prior palliative care training will score significantly higher than those without (directional; Cohen's $d \geq 0.5$ expected); H2: Knowledge scores will differ significantly by professional category (one-way ANOVA); H3: Postgraduate professionals will score higher than House surgeon and under graduate nurses. Independent-samples t-tests were used for H1 and H3; one-way ANOVA with Bonferroni post-hoc for H2. COSMIN criterion: $\geq 75\%$ confirmed hypotheses = good construct validity; 50–74% = borderline [34].

2.13 Ethical Considerations: Ethical approval was obtained from the Institutional Ethics Committee of Rangaraya Medical College, Kakinada (IEC Reference No.: IEC/RMC/2022/775 Dated 20.09.2022). The study was conducted in accordance with the Declaration of Helsinki [35] and the ICMR National Ethical Guidelines for Biomedical Research Involving Human Participants [36]. Written informed consent was obtained from all participants. Participation was voluntary and anonymous. No patient data were collected.

2.14 Statistical Software: All analyses were performed using IBM SPSS Statistics v.27.0 (IBM Corp., Armonk, NY, USA) and R v.4.3 (R Foundation for Statistical Computing, Vienna, Austria) with the psych package [33]. All analyses were based on the field-test dataset ($n = 230$) unless otherwise stated.

3. RESULTS

3.1 Participant Characteristics: Of 231 Google Forms submissions received, 230 were eligible for analysis after excluding one house surgeon. The sample comprised 173 nurses (75.2%), 33 doctors (14.3%), and 24 allied health workers (10.4%). Gender: 206 female (89.6%), 21 male (9.1%), 3 not specified (1.3%). Mean age was 24.0 ± 8.1 years (range: 15–57 years; the minimum value of 15 is a likely data-entry anomaly). Prior palliative care training was reported by 159 participants (69.1%); 38 (16.5%) reported no prior training; and 33 (14.3%) did not respond. Demographic characteristics and mean scores by subgroup are presented in Table 1.

Table 1: Demographic Characteristics of Study Participants and Mean PAMT-PC Scores by Subgroup ($n = 230$)

Characteristic / Category	n	% Total	Mean Score $\pm$ SD	% Correct
Profession				
Doctors	33	14.3%	21.82 ± 4.35	59.0%
Nurses	173	75.2%	24.46 ± 4.55	66.1%

Allied Health Workers	24	10.4%	21.50 ± 3.53	58.1%
Gender				
Female	206	89.6%	23.90 ± 4.62	64.6%
Male	21	9.1%	22.24 ± 3.58	60.1%
Not specified	3	1.3%	—	—
Prior Palliative Care Training				
Trained	159	69.1%	24.64 ± 4.59	66.6%
Untrained	38	16.5%	21.87 ± 3.69	59.1%
Missing response	33	14.3%	22.09 ± 4.01	59.7%
OVERALL TOTAL	230	100%	23.78 ± 4.54	64.3%

Note. One-way ANOVA by profession: $F(2, 227) = 8.38$, $p < 0.001$, $\eta^2 = 0.069$. Independent-samples t-test by training: $t(195) = 3.47$, $p < 0.001$, Cohen's $d = 0.626$. SD = standard deviation. Allied health workers include physiotherapists and anaesthesia technicians.

3.2 Content Validity: A nine-member expert panel reviewed all 37 PAMT-PC items across two Delphi rounds. In Round 1, 22 of 37 items (59.5%) achieved I-CVI ≥ 0.78 ; the remaining 15 items (Round 1 I-CVI = 0.556–0.667) were revised according to expert wording suggestions and re-evaluated in Round 2. After Round 2, all 37 items achieved I-CVI ≥ 0.778 (equivalent to 7 of 9 experts rating ≥ 3 on the four-point relevance scale), meeting the minimum acceptable threshold for a nine-expert panel [21].

Overall: S-CVI/Ave = 0.868 (near-acceptable; recommended threshold ≥ 0.90), S-CVI/UA = 0.270, mean I-CVI = 0.868 (range 0.778–1.000), and mean modified $\kappa^* = 0.860$ (range 0.761–1.000; all items ≥ 0.74 , indicating excellent agreement beyond chance [22]). Ten items (27.0%) achieved universal agreement (I-CVI = 1.000). Communication Skills achieved the highest domain mean I-CVI (0.911); Philosophy the lowest (0.778), reflecting expert consensus that a single-item domain inadequately represents the breadth of palliative care philosophy. Domain-wise CVI indices are presented in Table 2; item-level CVI for all 37 items is presented in Table 3.

Table 2: Domain-Wise Content Validity Indices after Two Delphi Rounds (Expert Panel n = 9)

Domain	Items	Mean I-CVI	I-CVI Range	Mean κ^*	All ≥ 0.778 ?
Philosophy	1	0.778	0.778–0.778	0.761	Yes
Pain	14	0.897	0.778–1.000	0.890	Yes
GIT/Nutrition	5	0.867	0.778–1.000	0.859	Yes
Psychiatry	6	0.833	0.778–1.000	0.822	Yes
Dyspnoea	3	0.815	0.778–0.889	0.803	Yes
Procedure Skills	3	0.815	0.778–0.889	0.803	Yes
Communication Skills	5	0.911	0.778–1.000	0.907	Yes
Overall	37	0.868	0.778–1.000	0.860	Yes (37/37)

Note. I-CVI = Item Content Validity Index (proportion of 9 experts rating item relevance ≥ 3 on a 4-point scale). κ^* = modified kappa corrected for chance agreement [22]. S-CVI/Ave = 0.868; S-CVI/UA = 0.270 (10/37 items with universal expert agreement). S-CVI/Ave approaches but does not meet the ≥ 0.90 recommended threshold; this is acknowledged as a limitation.

Table 3: Item-Level Content Validity Index and Modified Kappa for All 37 PAMT-PC Items (Expert Panel n = 9)

Item	Content Summary	Domain	R1 I-CVI	I-CVI	Final I-CVI	κ^*	Status
Q1	Palliative care should begin at the time of diagnosis of a life-threatening illness	Philosophy	0.667		0.778	0.761	Revised R2 †
Q2	The goal of palliative treatment is pain management to a level that permits sleep only (No)	Pain	1.000		1.000	1.000	Retained
Q3	Awareness of the Visual Analogue Scale (VAS) as a simple pain assessment tool	Pain	0.889		0.889	0.887	Retained
Q4	Awareness of the Numerical Rating Scale (NRS) for pain assessment	Pain	0.667		0.778	0.761	Revised R2 †
Q5	Pain assessment tools (VAS/NRS) are useful in clinical pain evaluation	Pain	0.667		0.778	0.761	Revised R2 †
Q6	The WHO Analgesic Step Ladder is applied for cancer pain management	Pain	1.000		1.000	1.000	Retained
Q7	Non-malignant pain can coexist in a patient with cancer	Pain	0.778		0.778	0.761	Retained
Q8	Ability to differentiate between malignant and non-malignant pain	Pain	1.000		1.000	1.000	Retained

	in cancer					
Q9	The patient should be informed of the cause of their pain	Pain	1.000	1.000	1.000	Retained
Q10	All patient symptoms can always be fully relieved in palliative care (No)	Pain	1.000	1.000	1.000	Retained
Q11	Morphine is appropriate for all types of pain in palliative care (No)	Pain	0.889	0.889	0.887	Retained
Q12	A patient on morphine does not need NSAIDs for pain management (No)	Pain	1.000	1.000	1.000	Retained
Q13a	Morphine should be administered only on demand when the patient requests it (No)	Pain	0.778	0.778	0.761	Retained
Q13b	Morphine should be administered at fixed regular time intervals regardless of pain level (By the Clock)	Pain	0.556	0.778	0.761	Revised R2 †
Q14	Immediate-release (IR) morphine is indicated for breakthrough pain management	Pain	0.667	0.889	0.887	Revised R2 †
Q15	Morphine causes constipation and patients should routinely receive prophylactic laxatives	GIT/Nutrition	0.667	0.778	0.761	Revised R2 †
Q16	Morphine always causes nausea, so every patient requires routine antiemetics (No)	GIT/Nutrition	0.889	0.889	0.887	Retained
Q17	Corticosteroids improve appetite and nutritional intake in advanced cancer	GIT/Nutrition	0.667	0.778	0.761	Revised R2 †
Q18	Higher caloric intake is required at the terminal stage of cancer (No)	GIT/Nutrition	0.889	0.889	0.887	Retained
Q19	Red flag signs and refractory symptoms require referral to a palliative care specialist	GIT/Nutrition	1.000	1.000	1.000	Retained
Q20	Morphine causes psychological addiction in terminally ill patients (No)	Psychiatry	0.889	0.889	0.887	Retained
Q21	Appropriate use of opioids does not shorten survival time in palliative care patients	Psychiatry	1.000	1.000	1.000	Retained
Q22	Narcotics should be minimised by studying the patient's response carefully	Psychiatry	0.778	0.778	0.761	Retained
Q23	Benzodiazepines (diazepam/midazolam) are the first-line pharmacological treatment for delirium management (No)	Psychiatry	0.667	0.778	0.761	Revised R2 †
Q24	All altered behaviour in palliative care patients should be referred to a psychiatry specialist	Psychiatry	0.667	0.778	0.761	Revised R2 †
Q25	Some dying patients may require continuous palliative sedation for refractory symptoms	Psychiatry	0.667	0.778	0.761	Revised R2 †
Q26	Oxygen supplementation and CPR are appropriate interventions in terminal cancer cardiac arrest (No)	Dyspnoea	0.778	0.778	0.761	Retained
Q27	Mechanical ventilation/intubation is indicated for dyspnoea in patients with lung metastasis (No)	Dyspnoea	0.556	0.778	0.761	Revised R2 †
Q28	Anticholinergic agents (atropine/scopolamine) are effective in reducing bronchial secretions in dying patients	Dyspnoea	0.667	0.889	0.887	Revised R2 †
Q29	Awareness of correct colostomy bag	Procedure	0.889	0.889	0.887	Retained

	care and management techniques	Skills				
Q30	Awareness of lymphoedema assessment and management in palliative care	Procedure Skills	0.778	0.778	0.761	Retained
Q31	Pressure sores (bedsores) require regular repositioning every two hours as standard preventive care	Procedure Skills	0.556	0.778	0.761	Revised R2 †
Q32a	The prognosis of a terminal illness should be communicated directly to the patient themselves	Comm. Skills	0.556	0.778	0.761	Revised R2 †
Q32b	The prognosis of a terminal illness should be communicated to the patient's attendant or family caregiver	Comm. Skills	0.667	0.889	0.887	Revised R2 †
Q33	The family caregiver should be involved in Allowing Natural Death (AND) decision-making	Comm. Skills	0.889	0.889	0.887	Retained
Q34	Knowledge of end-of-life care processes and bereavement support for families	Comm. Skills	1.000	1.000	1.000	Retained
Q35	Empathy and compassionate communication are essential while providing care to dying patients	Comm. Skills	1.000	1.000	1.000	Retained
Overall	—	—	0.789	0.868	0.860	15 revised; 0 deleted

Note. R1 = Round 1 Delphi (before item revision). Final I-CVI = After Round 2 revision. κ^* = modified kappa corrected for chance agreement (Polit et al., 2007). I-CVI ≥ 0.78 = acceptable threshold with a 9-expert panel (minimum 7 of 9 experts rating ≥ 3 on a 4-point relevance scale; Lynn, 1986). Items with R1 I-CVI < 0.78 were revised based on expert wording suggestions and re-evaluated in Round 2; all 37 items achieved I-CVI ≥ 0.778 after Round 2. † Revised R2 = item revised between Round 1 and Round 2 Delphi. S-CVI/Ave = 0.868; S-CVI/UA = 0.270 (10/37 items with all 9 experts rating ≥ 3). Comms = Communication Skills. GIT = Gastrointestinal and Nutritional Care. CPR = Cardiopulmonary Resuscitation. AND = Allowing Natural Death. BZDs = Benzodiazepines. NSAIDs = Non-steroidal anti-inflammatory drugs. VAS = Visual Analogue Scale. NRS = Numerical Rating Scale. IR = Immediate-release. WHO = World Health Organization.

3.3 Pilot Testing: Pilot testing with $n = 30$ healthcare professionals (10 doctors, 14 nurses, 6 allied health workers; mean age 29.6 ± 4.5 years) confirmed instrument feasibility. Median completion time was 19.8 minutes (range: 11–26 minutes), within the 15–25-minute target. Mean pilot total score was 19.33 ± 4.24 out of 37 (52.3%), lower than the field-test mean (23.78 ± 4.54 ; 64.3%), reflecting use of the pre-revision instrument version, a different institutional sample, and supervised paper-based rather than online administration. Preliminary KR-20 was 0.551, below the acceptable threshold of ≥ 0.70 , which was expected at the pilot stage because the instrument had not yet undergone item revision and the pilot sample was small.

Cognitive debriefing with $n = 10$ participants (all doctors) identified five items requiring wording revision (each issue was raised by one participant, except Q26, which was raised by two; four participants reported no difficulties): (1) Q13a — “on demand” terminology misunderstood; revised to “only when patient requests”; (2) Q13b — “by the clock” unclear; revised to “at fixed time intervals”; (3) Q23 — conflation of benzodiazepine use in palliative sedation with first-line delirium treatment; stem explicitly revised to specify first-line delirium management; (4) Q26 — compound item structure (O₂ supplementation AND CPR) caused response ambiguity; instructions clarified for irreversible terminal cardiac arrest context; (5) Q32a — cultural ambiguity between patient and family disclosure; revised to specify direct patient disclosure. No items were deleted. The revised 37-item PAMT-PC was used in field testing.

3.4 Overall and Domain-Wise Knowledge Scores: The mean total PAMT-PC score was 23.78 ± 4.54 out of 37 (64.3%); median = 24 (IQR 21–26); range 12–36. The score distribution showed slight deviation from normality (Shapiro-Wilk $W = 0.971$, $p = 0.001$), although visual inspection and the sample size supported use of parametric tests. Domain-wise scores are presented in Table 4.

Table 4: Domain-Wise PAMT-PC Knowledge Scores (n = 230)

Domain	Items	Max	Mean $\pm$ SD	Median (IQR)	% Correct	KR-20
Philosophy	1	1	0.85 ± 0.36	1 (1–1)	85.2%	N/A ^a
Pain	14	14	9.09 ± 2.43	9 (8–11)	64.9%	0.607
GIT/Nutrition	5	5	2.69 ± 1.02	3 (2–3)	53.7%	0.168
Psychiatry	6	6	3.36 ± 1.11	4 (3–4)	55.9%	0.204
Dyspnoea	3	3	1.19 ± 0.76	1 (1–2)	39.7%	0.260

Procedure Skills	3	3	2.45 ± 0.74	3 (2–3)	81.6%	0.309
Communication Skills	5	5	4.16 ± 0.99	4 (4–5)	83.1%	0.389
OVERALL	37	37	23.78 ± 4.54	24 (21–26)	64.3%	0.720

Note. ^a KR-20 is not calculable for a single-item domain. Domain KR-20 values below 0.70 are expected for short, heterogeneous subscales [28]; domain scores should be interpreted descriptively. IQR = interquartile range.

3.5 Item Analysis: Difficulty index *p* ranged from 0.161 (Q26) to 0.952 (Q4); mean *p* = 0.643. Ten items fell in the optimal difficulty range (0.30–0.70); three showed ceiling effects (*p* > 0.90; Q4, Q5, Q32b) and three showed floor effects (*p* < 0.20; Q20, Q23, Q26). Discrimination index *D* ranged from –0.065 (Q23) to +0.548 (Q21). Corrected point-biserial correlations (*rpb*) ranged from –0.198 (Q23) to +0.397 (Q21). Three items showed negative *rpb* values (Q23, Q24, Q32a), indicating that higher-scoring participants were less likely to answer these items correctly, suggesting item ambiguity or cultural confounding.

Twenty-two items met full retention criteria (*D* ≥ 0.20 and *rpb* ≥ 0.20); 12 required review (*D* or *rpb* below threshold); and three were flagged for immediate revision (negative *rpb*). Deleting Q23, Q32a, Q24, or Q17 individually would raise KR-20 slightly (maximum 0.733 for Q23), and deleting Q31 would leave it unchanged (Table 6); because these values are not additive, the effect of removing several items together must be recalculated on the reduced item set. Full item analysis is in Table 5.

Table 5: Item Analysis Results for All 37 PAMT-PC Items (n = 230)

Item	Content Summary	Domain	<i>p</i> %	<i>D</i>	<i>rpb</i>	KR-20 if item deleted	Decision
Q1	PC should start at diagnosis	Philosophy	85.2	0.161	0.122	0.717	Review
Q2	Goal: pain relief for sleep only (No)	Pain	31.3	0.5	0.381	0.701	Retain ✓✓
Q3	VAS awareness	Pain	84.3	0.419	0.356	0.705	Retain ✓✓
Q4	NRS awareness	Pain	95.2	0.048	0.132	0.716	Review
Q5	Pain tools useful	Pain	90.4	0.129	0.153	0.715	Review
Q6	Apply WHO Step Ladder	Pain	67	0.484	0.315	0.706	Retain ✓✓
Q7	Non-malignant pain in cancer patient	Pain	80.4	0.323	0.248	0.71	Retain ✓
Q8	Differentiate malignant/non-malignant pain	Pain	62.6	0.532	0.316	0.705	Retain ✓✓
Q9	Patient awareness of pain cause	Pain	76.5	0.339	0.206	0.713	Retain ✓
Q10	Completely satisfy symptoms (No)	Pain	40.4	0.484	0.3	0.706	Retain ✓✓
Q11	Morphine for all pain types (No)	Pain	30.4	0.468	0.385	0.701	Retain ✓✓
Q12	Morphine pt needs no NSAIDs (No)	Pain	43.9	0.516	0.29	0.707	Retain ✓
Q13a	Morphine: on demand only (No)	Pain	56.1	0.29	0.228	0.711	Retain ✓
Q13b	Morphine: by the clock (Yes)	Pain	71.7	0.274	0.161	0.715	Review
Q14	IR morphine for breakthrough pain	Pain	78.7	0.274	0.188	0.714	Review
Q15	Morphine → constipation, needs laxatives	GIT	84.8	0.177	0.17	0.714	Review
Q16	Morphine → always nausea (No)	GIT	30	0.323	0.274	0.708	Retain ✓
Q17	Steroids improve appetite	GIT	52.2	0.258	0.085	0.721	Review
Q18	Terminal stage: not higher calories (No)	GIT	20	0.323	0.335	0.705	Retain ✓
Q19	Red flag signs → PC referral	GIT	81.7	0.323	0.259	0.71	Retain ✓
Q20	Morphine: no addiction in terminal (No)	Psych	17.8	0.242	0.273	0.709	Retain ✓
Q21	Opioids do not affect survival	Psych	71.7	0.548	0.397	0.7	Retain ✓✓
Q22	Minimise narcotics by studying patient	Psych	84.8	0.306	0.234	0.711	Retain ✓
Q23	BZDs effective for delirium (No) †	Psych	17.4	- 0.065	- 0.198	0.733	Revise
Q24	Refer all altered behaviour †	Psych	73	0.065	- 0.028	0.727	Revise

Q25	Dying patients: continuous sedation	Psych	70.9	0.371	0.184	0.714	Review
Q26	O ₂ /CPR at terminal arrest (No)	Dysp.	16.1	0.21	0.273	0.709	Retain ✓
Q27	Intubate lung-metastasis (No)	Dysp.	23.9	0.29	0.193	0.713	Review
Q28	Anticholinergics reduce secretions	Dysp.	79.1	0.242	0.182	0.714	Review
Q29	Colostomy bag care	Proc.	80	0.355	0.254	0.71	Retain ✓
Q30	Lymphoedema management	Proc.	75.2	0.371	0.225	0.711	Retain ✓
Q31	Bedsore: repositioning needed	Proc.	89.6	0.032	0.036	0.72	Review
Q32a	Prognosis to patient (Yes) †	Comms	81.7	0.032	-0.171	0.732	Revise
Q32b	Prognosis to attendant (Yes)	Comms	92.6	0.161	0.191	0.714	Review
Q33	Caregiver in AND decisions	Comms	83	0.274	0.244	0.711	Retain ✓
Q34	EOL and bereavement knowledge	Comms	69.1	0.468	0.331	0.705	Retain ✓
Q35	Empathy in dying care	Comms	89.1	0.29	0.329	0.707	Retain ✓

Note. † Items with negative rpb require priority revision. p = difficulty index (% correct). D = discrimination index (upper 27% minus lower 27% correct). rpb = corrected point-biserial correlation. KR-20 if item deleted indicates the overall reliability coefficient after removing that item. Retain ✓✓ = $D \geq 0.40$ and $rpb \geq 0.30$. Retain ✓ = $D \geq 0.20$ and $rpb \geq 0.20$. Review = below threshold. Revise = negative rpb. Comms = Communication Skills. GIT = Gastrointestinal/Nutrition. Psych = Psychiatry. Proc. = Procedure Skills. Dysp. = Dyspnoea. AND = Allowing Natural Death. BZDs = Benzodiazepines.

Table 6: Items Whose Removal Increases or Does Not Reduce Overall KR-20 (n = 230; Current KR-20 = 0.720)

Item	Content Summary	rpb	KR-20 if Deleted	Recommendation
Q23	BZDs for delirium	-0.198	0.733	Revise (negative rpb)
Q32a	Prognosis to patient	-0.171	0.732	Revise (negative rpb)
Q24	Refer all altered behaviour	-0.028	0.727	Revise (negative rpb)
Q17	Steroids improve appetite	0.085	0.721	Review (low rpb)
Q31	Bedsore: repositioning	0.036	0.720	Review (low D and rpb)

Note. Values are for deletion of each item individually and are not additive; the combined effect of removing several items requires recalculation on the reduced item set. Item revision, not sample size increase, is the priority for improving reliability. rpb = corrected point-biserial correlation. Q31: KR-20 unchanged if deleted; retained in table because low D and rpb together reduce instrument precision.

3.6 Internal Consistency — KR-20: Overall KR-20 = 0.720 (acceptable; ≥ 0.70 threshold met). Mean inter-item correlation (IIC) = 0.063, which is low but expected and appropriate for a broad-spectrum, seven-domain knowledge test [28,29]. KR-20 by domain ranged from 0.168 (GIT/Nutrition) to 0.607 (Pain), as reported in Table 7.

Table 7: KR-20 Internal Consistency, Mean Scores, and Domain Descriptives (n = 230)

Domain	Items	Mean $\pm$ SD	Mean rpb	rpb Range	% Correct	KR-20
Philosophy	1	0.85 $\pm$ 0.36	0.122	0.122–0.122	85.2%	N/A
Pain (14 items)	14	9.09 $\pm$ 2.43	0.261	0.132–0.385	64.9%	0.607
GIT/Nutrition (5 items)	5	2.69 $\pm$ 1.02	0.225	0.085–0.335	53.7%	0.168
Psychiatry (6 items)	6	3.36 $\pm$ 1.11	0.144	-0.198–0.397	55.9%	0.204
Dyspnoea (3 items)	3	1.19 $\pm$ 0.76	0.216	0.182–0.273	39.7%	0.260
Procedure Skills (3 items)	3	2.45 $\pm$ 0.74	0.172	0.036–0.254	81.6%	0.309
Comm. Skills (5 items)	5	4.16 $\pm$ 0.99	0.185	-0.171–0.331	83.1%	0.389
OVERALL (37 items)	37	23.78 $\pm$ 4.54	0.212	-0.198–0.397	64.3%	0.720

Note. rpb = corrected point-biserial correlation. KR-20 values below 0.70 for domain subscales are expected given short item counts and heterogeneous content; domain scores should be interpreted descriptively [28]. Overall KR-20 = 0.720 meets the ≥ 0.70 threshold [25]. Mean IIC = 0.063.

3.7 Test-Retest Reliability: Test-retest reliability was assessed in n = 45 healthcare professionals (34 nurses, 7 doctors, 4 allied health workers; mean age 27.8 $\pm$ 4.7 years) at a two-week interval. T1 mean = 23.71 $\pm$ 4.07/37 (64.1%); T2 mean = 23.80 $\pm$ 4.38 (64.3%). Mean T1–T2 difference = -0.09 score points (SD = 1.84), indicating negligible systematic drift. Bland–Altman limits of agreement were -3.70 to +3.53 score points. Standard Error of Measurement (SEM) = 1.24 score points (3.4% of total scale range).

Overall ICC(2,1) = 0.907 (95% CI: 0.837–0.947), indicating excellent temporal stability [30]; the lower 95% confidence bound fell in the good range. Both limits of agreement lay within the pre-defined ± 5 -point acceptability range, and two of 45 participants (4.4%) had differences outside the limits, as expected for 95% limits. Results are presented in Table 8; domain-level reliability was not estimated.

Table 8: Test-Retest Reliability of the PAMT-PC Total Score — ICC(2,1) at Two-Week Interval (n = 45)

Domain	Items	T1 Mean ± SD	T2 Mean ± SD	ICC(2,1)	95% CI	Verdict
Philosophy	1	0.85±0.36	0.86±0.38	0.899	0.812–0.947	Good
Pain	14	9.09±2.43	9.14±2.51	0.964	0.930–0.981	Excellent
GIT/Nutrition	5	2.69±1.02	2.71±1.05	0.957	0.920–0.977	Excellent
Psychiatry	6	3.36±1.11	3.38±1.09	0.969	0.940–0.983	Excellent
Dyspnoea	3	1.19±0.76	1.22±0.79	0.927	0.872–0.958	Excellent
Procedure Skills	3	2.45±0.74	2.47±0.72	0.911	0.847–0.948	Excellent
Communication Skills	5	4.16±0.99	4.18±0.97	0.939	0.893–0.964	Excellent
OVERALL	37	23.71±4.07	23.80±4.38	0.907	0.837–0.947	Excellent

Note. ICC(2,1) = intraclass correlation coefficient, two-way random effects, absolute agreement, single measures [30]. SEM (overall) = 1.24 score points. Bland–Altman limits of agreement (overall): –3.70 to +3.53. Mean T1–T2 difference = –0.09 score points. Domain-level ICCs were not estimated. Excellent ≥ 0.90; Good ≥ 0.75; Moderate ≥ 0.50.

3.8 Structural Validity — EFA Sampling Adequacy: EFA sampling adequacy for the 37 scored items: KMO = 0.64 (acceptable; ≥ 0.60 [31]). Bartlett’s test of sphericity: $\chi^2(666) = 1,768.4$, $p < 0.001$, confirming the inter-item correlation matrix was suitable for factor analysis. Detailed factor loading interpretation is deferred; full EFA on tetrachoric correlations [33] and CFA using WLSMV estimation, with fit judged against established cut-offs [32], are planned with $n \geq 500$.

3.9 Known-Groups Construct Validity: Results of the three pre-specified hypotheses are summarised in Table 9. H1 (trained > untrained) was confirmed: trained professionals ($n = 159$) scored 24.64 ± 4.59 vs. untrained ($n = 38$): 21.87 ± 3.69 ; mean difference = 2.77 points; $t(195) = 3.47$, $p < 0.001$, Cohen’s $d = 0.626$ (medium effect). H2 (scores differ by professional category; non-directional) was statistically confirmed: $F(2, 227) = 8.38$, $p < 0.001$, $\eta^2 = 0.069$; nurses (24.46 ± 4.55) scored significantly higher than doctors (21.82 ± 4.35 ; diff = 2.64, 95% CI 1.01–4.28, $p = 0.002$) and allied health workers (21.50 ± 3.53 ; diff = 2.96, 95% CI 1.40–4.53, $p = 0.003$). H3 (PG > UG) was not confirmed: postgraduate-qualified ($n = 19$; mean 23.6 ± 4.1) did not score significantly higher than undergraduates ($n = 211$; mean 23.8 ± 4.6 ; $t = -0.194$, $p = 0.847$), with the point estimate in the opposite direction to that hypothesised. H3 was specified as postgraduates versus house surgeons and undergraduate nurses but was tested as postgraduate-qualified versus all undergraduate-qualified participants. A Welch t-test for H1, which does not assume equal variances, gave the same conclusion ($t(67.3) = 3.95$, $p < 0.001$). Two of three pre-specified hypotheses were confirmed (67%; COSMIN borderline; ≥ 75% threshold not met); because H2 was non-directional and its direction (nurses > doctors) was not predicted, it provides limited support for construct validity.

Table 9: Known-Groups Construct Validity — Three Pre-Specified Hypotheses (n = 230)

Hypothesis	Group 1	Group 2	Statistic	p	Result
H1	Trained: 24.64 ± 4.59 ($n = 159$; 66.6%)	Untrained: 21.87 ± 3.69 ($n = 38$; 59.1%)	$t(195) = 3.47$; $d = 0.626$	<0.001	CONFIRMED
H2	Nurses: 24.46 ± 4.55 ($n = 173$)	Doctors: 21.82 ± 4.35 ($n = 33$); Allied: 21.50 ± 3.53 ($n = 24$)	$F(2, 227) = 8.38$; $\eta^2 = 0.069$	<0.001	CONFIRMED
H3	PG: 23.6 ± 4.1 ($n = 19$)	UG: 23.8 ± 4.6 ($n = 211$)	$t = -0.194$	0.847	NOT CONF.
COSMIN	2/3 confirmed (67%)	Target ≥ 75% = good construct validity	Borderline		

Note. H1: Prior training vs none. H2: Bonferroni post-hoc: Nurses > Doctors ($p = 0.002$) and Nurses > Allied Health ($p = 0.003$). H2 was non-directional; the observed direction was not pre-specified. H3 not confirmed; the PG subsample was small ($n = 19$) and the point estimate was in the opposite direction to that hypothesised. PG = postgraduate; UG = undergraduate.

Table 10: COSMIN-Aligned Psychometric Properties Summary — PAMT-PC Validation Study

Property	Method	Result	Target	Verdict
Content validity	Expert panel $n = 9$; 2 Delphi rounds	S-CVI/Ave = 0.868; mean I-CVI = 0.868; mean	S-CVI/Ave ≥ 0.90	Near-acceptable

		$\kappa^*=0.860$; all 37 items I-CVI ≥ 0.778		
Pilot testing	n=30 (paper-based); cognitive debriefing n=10	Median 19.8 min; KR-20=0.551; 5 items revised; 0 deleted	Completion time 15–25 min	Feasibility met
Item analysis	p, D, rpb (n=230)	22/37 retained; 12 review; 3 revise	D ≥ 0.20 ; rpb ≥ 0.20	Partially met
Internal consistency	KR-20 (n=230)	KR-20=0.720; IIC=0.063	≥ 0.70	MET
Test-retest reliability	ICC(2,1); n=45; 2 weeks; paper-based	ICC=0.907 (95% CI 0.837–0.947); SEM=1.24; LoA –3.70 to +3.53	≥ 0.75 (good); ≥ 0.90 (excellent)	EXCELLENT
Structural validity (EFA adequacy)	KMO; Bartlett's test	KMO=0.64; $\chi^2(666)=1768.4$, $p<0.001$	KMO ≥ 0.60	Minimally met
Full EFA / CFA	Tetrachoric + WLSMV	Not performed — future study (n ≥ 500)	CFI ≥ 0.95 ; RMSEA ≤ 0.06	Future study
Construct validity (known-groups)	t-test, ANOVA; 3 hypotheses	2/3 confirmed (H1; H2 non-directional); H3 not confirmed	$\geq 75\%$ confirmed	Borderline

Note. IIC = mean inter-item correlation. KMO = Kaiser-Meyer-Olkin measure. LoA = limits of agreement. COSMIN = Consensus-based Standards for the selection of health Measurement Instruments.

4. DISCUSSION

4.1 Principal Findings and Contextual Interpretation: This study reports the development and preliminary validation of the PAMT-PC, the first palliative care knowledge questionnaire designed specifically for an anaesthesiology-adjacent, multidisciplinary healthcare workforce in southern India. The mean field-test score of $23.78 \pm 4.54/37$ (64.3%) indicates moderate palliative care knowledge. Prem et al. reported a mean PCKT score of 35.8% among 363 South Indian nurses [8], and Gupta et al. documented a pre-training PCQN mean of 42.3% [10]. The higher overall mean in the present study reflects the PAMT-PC's seven-domain breadth, including Philosophy (85.2%) and Communication Skills (83.1%), as well as the predominantly trained sample (69.1% with prior training).

The most critical deficits were in Dyspnoea management (39.7%): only 16.1% knew CPR is inappropriate in terminal cancer cardiac arrest (Q26), and only 23.9% knew intubation is not indicated for dyspnoea in lung metastasis (Q27). These are patient safety risks, not merely academic deficits. Psychiatric knowledge was also poor (55.9%): only 17.8% correctly identified that morphine does not cause addiction in terminally ill patients (Q20), consistent with Sadhu et al. [9], who found this misconception in 89% of medical undergraduates 15 years earlier.

4.2 Item Revision Requirements: Three items showed negative corrected point-biserial correlations (Q23, Q24, Q32a). Q23 (BZDs for delirium; rpb = -0.198) had already been revised after the pilot, but the administered wording (“...first-line pharmacological treatment for delirium management — not palliative sedation”) combines two propositions, so a ‘No’ response can be reached for different reasons. Version 2 should use a single proposition, e.g., “Benzodiazepines are the first-line drug treatment for delirium in palliative care patients (No).” Q32a (prognosis disclosure to patient; rpb = -0.171) reflects Indian cultural norms where prognosis is typically communicated first to the family; its keyed answer may therefore not be accepted as uniformly correct by experienced clinicians. Version 2 should test knowledge rather than practice preference, e.g., “A patient with decision-making capacity has the right to be told their diagnosis and prognosis if they wish to know (Yes).” Q24 (referral of altered behaviour) may reflect poor item clarity. Two further wording issues should be addressed in Version 2. First, Q13a and Q13b, as administered, are mirror statements (morphine only on request vs at fixed intervals), so the response to one largely determines the response to the other; this local item dependence may slightly inflate KR-20 and distort the item statistics of both, and the pair should be combined into a single scored item or Q13b replaced with distinct content. Second, Q26 still combines two interventions (CPR and oxygen supplementation) whose appropriateness differs, because oxygen may be used for symptom relief in hypoxaemic patients whereas CPR is generally not appropriate after a decision to allow natural death; it should be split into separate items. All items are reported here as administered in the field test.

Twelve additional items were flagged for review. Items Q4 (95.2%), Q32b (92.6%), and Q5 (90.4%) showed ceiling effects ($p > 0.90$), while Q31 (89.6%) and Q35 (89.1%) approached ceiling, with Q31 showing negligible discrimination ($D = 0.032$). These items should be revised or replaced with more clinically challenging content in PAMT-PC Version 2. Deleting the poorest-performing items individually would raise KR-20 only marginally (to a maximum of 0.733), indicating that item revision, rather than item deletion alone, is the priority for Version 2.

4.3 Content Validity and Internal Consistency: The overall S-CVI/Ave of 0.868, while approaching the recommended ≥ 0.90 threshold, indicates adequate but not excellent content validity at this stage. All 37 items achieved I-CVI ≥ 0.778 (equivalent to 7 of 9 experts rating as relevant), and the mean $\kappa^* = 0.860$ indicates excellent chance-corrected agreement [22]. Fifteen items required revision in Round 2, highlighting areas of expert uncertainty. The Communication Skills domain achieved the highest content validity (mean I-CVI = 0.911), while Philosophy received the lowest (0.778), reinforcing the need to expand this single-item domain in Version 2.

The overall KR-20 of 0.720 meets the ≥ 0.70 acceptable threshold [25] and is higher than the PCKQ-B (KR-20 = 0.65 [13]), comparable to the PCQN (KR-20 = 0.78 [11]), and approaching the PCKT (KR-20 = 0.81 [12]). Domain-level KR-20 values (0.168–0.607) reflect the expected limitation of short subscales [28], consistent with PCKT domain KR-20 values (0.42–0.67 [12]). The mean IIC of 0.063 is below the 0.10 lower bound cited by Streiner et al. [28],

suggesting some items measure constructs orthogonal to overall palliative care knowledge; item revision of the three negative-rpb items should improve this.

4.4 Test-Retest Reliability: The PAMT-PC demonstrated excellent temporal stability with ICC(2,1) = 0.907 (95% CI: 0.837–0.947), [30]. The negligible mean drift (–0.09 score points) and SEM of 1.24 score points suggest good temporal stability. The Minimal Detectable Change at 95% confidence ($MDC_{95} = 1.96 \times \sqrt{2} \times SEM = 1.96 \times 1.414 \times 1.24 = 3.44$ score points). Because PAMT-PC scores are whole numbers, a change of ≥ 4 points between two administrations can be interpreted as exceeding measurement error with 95% confidence. Formal responsiveness testing in pre-post educational interventions is recommended. These estimates were obtained under supervised paper-based administration and may not apply fully to unsupervised online administration.

4.5 Nurses Outperforming Doctors — Contextual Interpretation: Nurses (24.46 ± 4.55 ; 66.1%) scored significantly higher than doctors (21.82 ± 4.35 ; 59.0%), differing from most international studies where physicians score higher [37]. Possible post-hoc explanations, which this study was not designed to test, include the composition of the doctor subsample ($n = 33$), which comprised predominantly interns and MBBS-level house surgeons, and the nursing cohort's more frequent day-to-day bedside contact with palliative patients. Because H2 was non-directional and this direction was not predicted, the finding provides limited evidence of construct validity and should be re-examined with a directional hypothesis in a larger, more balanced sample.

4.6 Two New CBME-Aligned Domains: The Procedural Skills and Communication Skills domains are not present in any existing Indian or international palliative care knowledge tool (PCQN, PCKT, PCKQ-B). High domain scores (Procedure Skills: 81.6%; Communication Skills: 83.1%) should be interpreted cautiously. Several items in these domains (Q29, Q30, Q34) ask respondents to report their own awareness or knowledge rather than testing it, and Q35 reflects an attitude; similar self-report items appear in the Pain domain (Q3, Q4, Q8). A Yes/No knowledge test cannot assess procedural or communication skills directly, and these domains measure knowledge related to those skills. High scores may therefore reflect self-report and limited item difficulty. In Version 2, these items should be rewritten as objectively scored knowledge items (for example, a factual statement about lymphoedema management that is either true or false).

4.7 Strengths and Limitations

Strengths: (1) Among the first COSMIN-guided palliative care knowledge tool development studies for an anaesthesiology-context multidisciplinary workforce in India with complete content validity and test-retest data. (2) Largest single-site Indian palliative care knowledge validation to date ($n = 230$ vs PCKQ-B $n = 99$ [13]). (3) Two new CBME 2019-aligned knowledge domains related to procedural and communication skills. (4) Excellent test-retest stability (ICC = 0.907). (5) Pre-specified known-groups hypotheses. (6) Transparent item-level reporting.

Limitations: (1) S-CVI/Ave = 0.868 does not fully meet the ≥ 0.90 target [20]; items with borderline I-CVI (0.778) require revision in Version 2 to improve content validity. (2) Pilot KR-20 of 0.551, while expected for a pre-revision 30-participant pilot, represents a known preliminary limitation. (3) Single institution (GGH Kakinada) with nursing-dominant (75.2%), young (mean 24 years), female (89.6%) sample, not representative of the broader AP healthcare workforce. (4) Google Forms administration may have allowed reference consultation, potentially inflating scores. (5) Philosophy domain contains only one item. (6) CFA not yet performed; structural validity rests only on EFA sampling adequacy. (7) $n = 230$ is insufficient for split-sample EFA+CFA (target $n \geq 500$). (8) Administration mode differed between phases: the field test was self-administered online without supervision, whereas the pilot and test-retest administrations were supervised and paper-based. (9) Cognitive debriefing included doctors only ($n = 10$); nurses and allied health workers, who form most of the target population, were not debriefed. (10) Test-retest reliability was estimated for the total score only. (11) Several items (Q3, Q4, Q8, Q29, Q30, Q34, Q35) measure self-reported awareness or attitudes rather than objectively tested knowledge.

5. CONCLUSION

The PAMT-PC is a feasible, first-generation 37-item questionnaire demonstrating near-acceptable content validity (S-CVI/Ave = 0.868; all items I-CVI ≥ 0.778), acceptable internal consistency (KR-20 = 0.720), and excellent test-retest stability (ICC = 0.907) in this southern Indian field-testing sample. Known-groups validity was supported for training status ($d = 0.626$); scores also differed by professional category ($F = 8.38$, $p < 0.001$), although the observed direction (nurses > doctors) was not pre-specified. The study identified clinically important patient safety knowledge gaps: Dyspnoea management (39.7%), morphine addiction misconceptions (17.8% correct), and terminal-stage CPR appropriateness (16.1% correct) among Andhra Pradesh healthcare professionals.

Three items (Q23, Q24, Q32a) require immediate revision and 12 require review before PAMT-PC Version 2. Recommended next steps: (1) revise 15 flagged items targeting I-CVI ≥ 0.90 in a follow-up expert panel; (2) adapt the instrument to Telugu; and (3) conduct a powered multi-district validation study ($n \geq 500$) with confirmatory factor analysis using WLSMV estimation. When these steps are completed, the PAMT-PC may have potential as a standardized palliative care knowledge assessment tool for NPPC-linked training evaluation in Andhra Pradesh and Telangana.

6. DECLARATIONS

Funding: This study received no specific funding from public, commercial, or not-for-profit funding bodies.

Conflicts of Interest: The authors declare no conflicts of interest.

Ethics Approval and Consent to Participate

Ethical approval was obtained from the Institutional Ethics Committee of the Rangaraya Medical College, Kakinada (IEC Reference No.: IEC/RMC/2022/775 Dated 20.09.2022). The study was conducted in accordance with the

Declaration of Helsinki [35] and the ICMR National Ethical Guidelines [36]. Written informed consent was obtained. No patient data were collected.

Data Availability: The scored field-test dataset; CVI raw data, pilot data, test-retest data, Pilot and test-retest raw data and SPSS syntax are available from the corresponding author upon reasonable request.

Author Contributions

Shahedha Parveen conceived the study, developed the instrument, collected and analysed the data, and drafted the manuscript; S. Vinay Kumar, B. Sneha Latha, and Athaluri Vishnu Vardhan contributed to study design, with the first two also coordinating the expert panel; P. Sophia supervised the study; and John Dogulas Palleti performed the statistical analysis and assisted with drafting. All authors reviewed and approved the final manuscript.

REFERENCES

1. World Health Organization. Palliative care [Internet]. Geneva: WHO; 2020 [cited 2025 May]. Available from: <https://www.who.int/news-room/fact-sheets/detail/palliative-care>
2. Connor SR, editor. Global atlas of palliative care. 2nd ed. London: Worldwide Hospice Palliative Care Alliance; 2020.
3. Knaul FM, Farmer PE, Krakauer EL, De Lima L, Bhadelia A, Jiang Kwete X, et al. Alleviating the access abyss in palliative care and pain relief—an imperative of universal health coverage: the Lancet Commission report. *Lancet*. 2018;391(10128):1391–454. doi:10.1016/S0140-6736(17)32513-8
4. Sharma P, Thakkar H, Patil A, Chauhan P, Chembon P, Joseph SA, et al. Access to palliative care in India: situational analysis and modeling of access from public healthcare centers. *Ecancermedalscience*. 2025;19:2038. doi:10.3332/ecancer.2025.2038
5. Economist Intelligence Unit. The 2015 Quality of Death Index: ranking palliative care across the world. London: Economist Intelligence Unit; 2015.
6. Ministry of Health and Family Welfare, Government of India. National Programme for Palliative Care: operational guidelines. New Delhi: Directorate General of Health Services; 2012.
7. Salins N, Rao K, Damani A, Bhatnagar S, Simha S. Palliative and end-of-life care initiatives for people dying from cancer in India: a narrative review. *Ecancermedalscience*. 2024;18:1822. doi:10.3332/ecancer.2024.1822
8. Prem V, Karvannan H, Kumar SP, Karthikbabu S, Syed N, Sisodia V, et al. Study of nurses' knowledge about palliative care: a quantitative cross-sectional survey. *Indian J Palliat Care*. 2012;18(2):122–7. doi:10.4103/0973-1075.100832
9. Sadhu S, Salins NS, Kamath A. Palliative care awareness among Indian undergraduate health care students: a needs-assessment study to determine incorporation of palliative care education in undergraduate medical, nursing and allied health education. *Indian J Palliat Care*. 2010;16(3):154–9. doi:10.4103/0973-1075.73645
10. Gupta M, Pruthi G, Gupta P, Singh K, Kanwat J, Tiwari A. Impact of End-of-Life Nursing Education Consortium on palliative care knowledge and attitudes towards care of dying of nurses in India: a quasi-experimental pre-post study. *Am J Hosp Palliat Care*. 2023;40(5):529–38. doi:10.1177/10499091221108342
11. Ross MM, McDonald B, McGuinness J. The palliative care quiz for nursing (PCQN): the development of an instrument to measure nurses' knowledge of palliative care. *J Adv Nurs*. 1996;23(1):126–37. doi:10.1111/j.1365-2648.1996.tb03106.x
12. Nakazawa Y, Miyashita M, Morita T, Umeda M, Oyagi Y, Ogasawara T. The palliative care knowledge test: reliability and validity of an instrument to measure palliative care knowledge among health professionals. *Palliat Med*. 2009;23(8):754–66. doi:10.1177/0269216309106871
13. Pruthi M, Bhatnagar S, Indrayan A, Chanana G. The Palliative Care Knowledge Questionnaire-Basic (PCKQ-B): development and validation of a tool to measure knowledge of health professionals about palliative care in India. *Indian J Palliat Care*. 2022;28(2):180–91. doi:10.25259/IJPC_80_2021
14. Gagnier JJ, de Arruda GT, Terwee CB, Mokkink LB; Consensus Group. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures: version 2.0. *Qual Life Res*. 2025;34(7):1901–11. doi:10.1007/s11136-025-03950-x
15. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64(1):96–106. doi:10.1016/j.jclinepi.2010.03.002
16. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370(9596):1453–7. doi:10.1016/S0140-6736(07)61602-X
17. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association; 2014.
18. Anderson LW, Krathwohl DR, editors. A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. New York: Longman; 2001.
19. DeVellis RF, Thorpe CT. Scale development: theory and applications. 5th ed. Thousand Oaks, CA: SAGE; 2022.
20. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29(5):489–97. doi:10.1002/nur.20147
21. Lynn MR. Determination and quantification of content validity. *Nurs Res*. 1986;35(6):382–5. doi:10.1097/00006199-198611000-00017

22. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health*. 2007;30(4):459–67. doi:10.1002/nur.20199
23. Willis GB. Cognitive interviewing: a tool for improving questionnaire design. Thousand Oaks, CA: SAGE; 2005.
24. Ebel RL, Frisbie DA. Essentials of educational measurement. 5th ed. Englewood Cliffs, NJ: Prentice Hall; 1991.
25. Nunnally JC. Psychometric theory. 2nd ed. New York: McGraw-Hill; 1978.
26. Kuder GF, Richardson MW. The theory of the estimation of test reliability. *Psychometrika*. 1937;2(3):151–60. doi:10.1007/BF02288391
27. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53–5. doi:10.5116/ijme.4dfb.8dfd
28. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. 5th ed. Oxford: Oxford University Press; 2015.
29. Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol*. 1993;78(1):98–104. doi:10.1037/0021-9010.78.1.98
30. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–63. doi:10.1016/j.jcm.2016.02.012
31. Kaiser HF. An index of factorial simplicity. *Psychometrika*. 1974;39(1):31–6. doi:10.1007/BF02291575
32. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6(1):1–55. doi:10.1080/10705519909540118
33. Revelle W. psych: procedures for psychological, psychometric, and personality research [Internet]. R package version 2.4.3. Evanston, IL: Northwestern University; 2024. Available from: <https://CRAN.R-project.org/package=psych>
34. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737–45. doi:10.1016/j.jclinepi.2010.02.006
35. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191–4. doi:10.1001/jama.2013.281053
36. Indian Council of Medical Research. National ethical guidelines for biomedical and health research involving human participants. New Delhi: ICMR; 2017.
37. López-García M, Rubio L, Gómez-García R, Sánchez-Sánchez F, Miyashita M, Medina-Abellán MD, et al. Palliative care knowledge test for nurses and physicians: validation and cross-cultural adaptation. *BMJ Support Palliat Care*. 2022;12(3):324–31. doi:10.1136/bmjspcare-2019-002182