Title: Design and Validation of a Health System Preparedness Assessment Tool for Mass-Gatherings

Authors: Fariba Ahmadi[1], Zohreh Ghomian[1,2,*], Sakinah Rakhshandehrou[3], Azadeh Tahernejad[4]

1. *Department of Health in Disasters and Emergencies, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran.*
2. *Safety Promotion and Injury Prevention Research Center, Institute for Health Sciences and Environment, Shahid Beheshti University of Medical Sciences, Tehran, Iran.*
3. *Department of Health in Education & Health Promotion, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran.*
4. *Department of Management Sciences and Health Economics, School of Health, Mashhad University of Medical Sciences, Mashhad, Iran.*

To appear in: ***Health in Emergencies & Disasters Quarterly***

Received date: 2025/01/22

Revised date: 2025/12/03

Accepted date: 2025/12/8

First Online Published: 2025/12/17

Please cite this article as:

Ahmadi F, Ghomian Z, Rakhshandehrou S, Tahernejad A. Design and Validation of a Health System Preparedness Assessment Tool for Mass-Gatherings. Health in Emergencies & Disasters Quarterly. Forthcoming 2026. Doi: http://dx.doi.org/10.32598/hdq.2026.671.1

Abstract

Background: Mass gatherings, such as the Arbaeen Pilgrims Walking, pose unique public health challenges requiring comprehensive preparedness strategies. Despite the existence of several tools, there is no validated and context-specific instrument to assess health system preparedness in such events.

Objective: This research aims to design and psychometrically validate the Health System Preparedness Assessment Tool for Mass Gatherings.

Methods: A sequential mixed-methods study was conducted in three phases. First, a scoping review of international and national documents was performed using Arksey and O'Malley's framework, resulting in the extraction of 324 unique preparedness-related items. In the second phase, qualitative expert consultations conducted through focus group discussions refined these items into four main domains and 11 subdomains. In the third phase, psychometric evaluation, including face, content, and construct validity, as well as reliability testing, was performed on data from 260 healthcare and emergency professionals.

Results: The final HPAT consisted of 148 items across four domains: Health, Medical Care and Treatment, Management and Logistics, and Coordination and Safety. Exploratory factor analysis confirmed the four-domain structure (KMO = 0.892; Bartlett's test $p < 0.001$). Content validity was excellent (mean CVI = 0.98), and internal consistency was also high (Cronbach's $\alpha = 0.96$). The tool demonstrated substantial inter-rater reliability (Kappa = 0.76).

Conclusion: The validated HPAT provides a reliable, comprehensive, and context-specific instrument for evaluating health system preparedness for mass gatherings. Its application can guide policy and practice by identifying gaps, strengthening intersectoral coordination, and enhancing public health readiness for large-scale events.

Keywords: Mass gatherings, Preparedness, Health system, Assessment.

1. Introduction

Mass gatherings, particularly those involving millions of participants, pose complex and multifaceted public health challenges. These events can lead to a wide spectrum of risks, including bioterrorism, terrorist attacks, natural disasters, fires, crowd crushes, and infectious disease outbreaks [1]. Globally, it is estimated that over the past three decades, more than 7,000 deaths and 14,000 injuries have occurred during mass gatherings. [2]. Health threats are often amplified by factors such as crowd density, emotional stress, and crowd dynamics, which can trigger secondary incidents [3]. Additionally, the nature of the gathering—whether cultural, political, or religious—can influence vulnerability levels and the specific health system demands [4].

Iran is one of the most prominent hosts of large-scale religious gatherings, with a calendar of 99 major political, cultural, and spiritual events annually [5]. Among these, the Arbaeen Pilgrims Walking (APW) stands out as one of the largest religious events in the world, drawing over 20 million participants each year [6]. The scale, cross-border nature, and unique cultural practices of the APW present significant challenges for health service delivery, including emergency response, disease prevention, crowd management, and cross-jurisdictional coordination [7].

Despite the recognized importance of preparedness for such high-risk gatherings, there is currently no comprehensive, standardized, and validated tool specifically designed to assess health system preparedness in culturally and religiously unique contexts such as the APW [8]. Existing tools are often designed for general public events, focusing primarily on broad domains such as emergency medical capacity or crowd control [9, 10]. They frequently overlook critical context-specific factors, including religious rituals, cultural norms, extreme environmental conditions, and psychosocial stressors that can influence health risks and service delivery [11, 12]. Moreover, most available instruments assess preparedness in a fragmented manner—examining single domains in isolation—rather than adopting a multidimensional approach that integrates personal, environmental, managerial, and public health perspectives into a single framework [13].

This research addresses these limitations by designing and validating the Health System Preparedness Assessment Tool for Mass Gatherings (HPAT), with an initial application to the APW context. The tool offers several unique advantages:

- Cultural specificity: Incorporates variables directly relevant to religious mass gatherings in Iran and similar settings, thereby ensuring contextual appropriateness [14].
- Multidimensional coverage: Simultaneously evaluates multiple domains—personal, environmental, managerial, and public health—providing a more holistic preparedness profile[15].
- Evidence-based development: Integrates data and insights from both qualitative and quantitative phases of previous research by the authors, combined with structured expert consensus sessions, ensuring both scientific rigor and practical relevance [16, 17].
- Psychometric validation: Undergoes rigorous testing for validity and reliability, addressing a key shortcoming of many existing checklists [18].

The proposed tool fills a critical methodological and operational gap by providing health systems with a reliable and culturally adapted instrument to assess readiness for mass gatherings. Its application can

help stakeholders identify vulnerabilities, allocate resources more effectively, and design targeted interventions that reduce human, financial, and social impacts. While this study focuses on the APW, the tool's adaptable structure enables it to be customized for other large-scale events worldwide, thereby contributing to the global mass-gathering health preparedness literature and practice.

2. Methods

This study employed a mixed-methods approach across three main phases: (1) a scoping review, (2) a qualitative component involving expert input, and (3) quantitative validation, with the ultimate aim of designing and psychometrically validating a checklist to assess health system preparedness during mass gatherings. Figure 1 illustrates the schematic overview of the research process.

2.1 The Review Phase

The first phase of the study involved a scoping review to identify existing tools, frameworks, and indicators relevant for assessing health system preparedness in the context of mass gatherings. This review followed the five-stage methodological framework proposed by Arksey and O'Malley, ensuring a systematic and replicable process.

Identifying the research question

The overarching question guiding the review was: "What tools, frameworks, or indicators have been developed to assess health system preparedness in mass gatherings?" This question was deliberately broad to capture both international and national evidence and to ensure comprehensive coverage of different types of preparedness measures.

Identifying relevant studies

A comprehensive search strategy was developed and applied to international and national scientific databases, including PubMed, Scopus, Web of Science, and Google Scholar, as well as Persian-language databases (SID, Magiran, and Irandoc). The search was conducted up to October 2023. Search terms combined controlled vocabulary and free-text keywords, including:

"mass gathering" AND "preparedness*" OR "readiness*" AND "tool" OR "instrument" OR "checklist."

Boolean operators were adapted to the indexing system of each database. Additionally, manual searches of reference lists and targeted searches of grey literature (e.g., national guidelines, policy documents, and unpublished reports) were conducted to ensure comprehensiveness.

Study selection

Inclusion criteria were: (a) studies, reports, or guidelines that described or evaluated preparedness assessment tools, frameworks, or checklists for health systems in mass gatherings; (b) documents published in English or Persian; and (c) materials accessible in full text. Exclusion criteria included publications focused solely on emergency preparedness unrelated to mass gatherings or studies without a clear methodological description.

Two independent reviewers screened titles and abstracts, followed by full-text screening of potentially eligible records. Any discrepancies were resolved through discussion, and, if necessary, through consultation with a third reviewer.

Charting the data

For each included document, data were systematically extracted using a structured form specifically developed by the research team. Extracted data included: study characteristics (author, year, country), type and purpose of tool/framework, domains covered (e.g., environmental, managerial, personal, or public health preparedness), and reported strengths and limitations. National guidelines and official documents were also mapped for their relevance and applicability to the Iranian context.

Collating, summarizing, and reporting the results

Extracted data were synthesized using thematic content analysis. Tools and frameworks were categorized according to their scope, domains, and methodological rigor. The synthesis highlighted common indicators across studies as well as gaps, particularly the absence of multidimensional tools tailored to religious and cultural contexts such as the Arbaeen Pilgrimage. These findings were integrated with the analysis of national documents and expert consultations to develop the preliminary item pool for the checklist.

2.2 The Qualitative Phase

The qualitative phase of this study aimed to contextualize, refine, and validate the initial framework of the HPAT. Data were collected through four structured expert discussion panels, each guided by the Nominal Group Technique (NGT). This approach was chosen to ensure systematic consensus-building and the inclusion of diverse perspectives.

Participants and Setting

A total of 12 experts participated, including health policymakers, academic researchers, and operational managers affiliated with the Ministry of Health and major medical universities in Iran. Purposive sampling was employed to ensure variation in professional backgrounds and expertise, enriching the diversity of perspectives. All participants had prior experience with planning, managing, or evaluating health services during mass gatherings, particularly large-scale religious events such as the Arbaeen Pilgrimage.

Data Collection

Four sessions were conducted between January and February 2024, each lasting approximately 90–120 minutes. The discussions were structured around the preliminary checklist items generated from the scoping review. The expert panels critically evaluated the practicality, cultural relevance, and contextual appropriateness of each item. Panelists also highlighted gaps, identified redundancies, and suggested modifications to ensure the checklist captured the complex realities of mass gatherings. Sessions were audio-recorded with informed consent, transcribed verbatim, and anonymized to maintain confidentiality.

Consensus was achieved through a structured voting process embedded within the NGT framework. Items were ranked and prioritized based on clarity, contextual fit, and comprehensiveness. This iterative process allowed the research team to refine the tool while retaining fidelity to both the scientific evidence and the lived expertise of practitioners and policymakers.

Data Analysis

Qualitative data were analyzed using content analysis. Transcripts were reviewed line by line to identify meaning units, which were then coded, categorized, and grouped into overarching themes relevant for refining checklist. Two independent researchers performed the initial coding, and any discrepancies were resolved through discussion and consensus. A third senior researcher reviewed the final thematic framework to ensure consistency of interpretation.

Trustworthiness

The rigor of the qualitative phase was ensured by following Lincoln and Guba's criteria for trustworthiness. Credibility was enhanced through triangulation of expert input with scoping review findings and through participant validation of the refined items. Transferability was supported by providing detailed descriptions of participant roles, the cultural context of the Arbaeen Pilgrimage, and the broader healthcare system setting. Dependability was maintained by keeping a transparent audit trail of analytic decisions, while confirmability was reinforced through reflexive discussions and peer debriefing within the research team.

Ethical Considerations

All experts provided written informed consent before participation. Discussions were anonymized to protect confidentiality.

2.3 Tool Development/Quantitative Phase

Integration of Findings

The development of the preliminary item pool followed a structured process that systematically combined the results of the scoping review and the qualitative phase. All potential indicators identified in the scoping review were initially collated and categorized. Next, insights from the expert panels using the Nominal Group Technique were incorporated to refine, contextualize, and expand the indicators, particularly for culturally specific settings such as the Arbaeen pilgrimage. Redundant or overlapping items were removed, while new suggestions were incorporated. The resulting draft checklist was organized into five domains: environmental

preparedness, personal preparedness, psychological readiness, health-related preparedness, and managerial preparedness. To ensure internal coherence, this draft underwent three independent rounds of review by the research team before proceeding to validation.

Face Validity

Face validity was assessed using both qualitative and quantitative approaches [19]. In the qualitative stage, twelve specialists in disaster health and emergency preparedness provided structured feedback on item clarity, comprehensiveness, and cultural applicability. Their feedback was incorporated into subsequent revisions. In the quantitative stage, the impact score was calculated for each item, based on its importance (5-point Likert scale) and frequency of endorsement [20]. Items with an impact score ≥1.5 were retained for further assessment [21, 22].

Content Validity

Content validity was assessed using both the Content Validity Ratio (CVR) and Content Validity Index (CVI) [23]. A total of 76 national experts participated in this phase. For CVR, the essentiality of each item was rated based on Lawshe's method, with a cut-off of ≥0.64 for retention. For CVI, experts rated relevance, clarity, and simplicity on a 4-point scale. Items with CVI ≥0.80 were accepted; items between 0.70–0.79 were revised, and those <0.70 were removed [24].

Construct Validity

Construct validity was assessed through a two-step process involving Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). The EFA was conducted with responses from 260 healthcare workers and emergency responders with prior experience in mass gatherings. Sampling adequacy was confirmed using the Kaiser-Meyer-Olkin (KMO) test (0.892) and Bartlett's test of sphericity ($p < 0.001$), demonstrating suitability for factor analysis. Factor extraction was performed using Principal Component Analysis with Varimax rotation. Items with factor loadings below 0.40 or substantial cross-loadings were reviewed and revised. The EFA results supported a four-dimensional framework of the tool, consistent with the conceptual model derived from the qualitative phase.

Subsequently, CFA was performed to test the adequacy of this four-factor model. Model fit was evaluated using multiple fit indices (χ^2/df, RMSEA, CFI, TLI), and standardized factor loadings were examined for significance and magnitude. This two-step approach provided robust evidence of the construct validity of the instrument.

Reliability

Reliability was examined using measures of internal consistency and inter-rater reliability. Internal consistency was assessed using Cronbach's alpha, with thresholds ≥0.80 considered acceptable. Inter-rater reliability was evaluated through the Kappa coefficient, where values

≥0.61 indicated substantial agreement. All statistical analyses were conducted using SPSS version 25 [15].

3. Results

3.1. Scoping Review

In the first phase, a total of 1,950 records were identified through systematic searches across international databases (PubMed, Scopus, Web of Science, and Google Scholar) and Persian-language databases (SID, Magiran, and Irandoc). After eliminating 1,608 duplicate records, 342 documents remained for title and abstract screening. Based on inclusion and exclusion criteria (e.g., relevance, language, and availability of full text), 272 records were excluded, leaving 70 full-text articles for detailed review. Following an in-depth evaluation, 9 studies were deemed relevant and included in the final scoping review.

In addition, five national policy documents developed over the last two decades, addressing emergency response, health surveillance, and disaster preparedness plans in Iran, were reviewed to ensure contextual comprehensiveness. From these, 193 initial codes were extracted. Combined with the 306 codes derived from the international articles, a total of 499 codes were initially obtained. After merging duplicate items and eliminating overlaps, 324 unique items were retained and classified into thematic domains.

The analysis revealed that preparedness for mass gatherings is multidimensional and requires integration of environmental, health system, managerial, psychosocial, and coordination-related aspects. The synthesized codes were grouped into five overarching themes and multiple subthemes. Table 2 presents a breakdown of article sources, the number of items extracted, and their classification into thematic domains. Table 3 summarizes the thematic classification of findings, showing both the scope and depth of the extracted evidence.

Table 1: Sources and Codes Extracted in the Scoping Review and Document Analysis

Source Type	Number of Documents	Codes/Items Extracted	Notes
International Articles	9	306	From databases: PubMed, Scopus, WoS, etc.
National Policy Documents	5	193	Official Iranian health preparedness docs
Total (before deduplication)	–	**499**	
Duplicates/Overlapping Items	–	175	Merged during synthesis
Final Unique Items	–	**324**	Used as an initial pool for checklist design

Table 2: Preliminary Domains and Subdomains Extracted from the Scoping Review and Document Analysis

Domain	Subdomains	Number of Items
Environmental Preparedness	Infrastructure, sanitation & crowd safety	65
Health-related Preparedness	Surveillance, essential services, and infection control	72
Managerial Preparedness	Governance, planning, training	61
Psychosocial Preparedness	Mental health, risk communication, and volunteers	54
Coordination & Collaboration	Intersectoral and international coordination	72
Total		324

3.2. Phase Two: Qualitative Validation and Expert Consultation

In the second phase, the 324 unique items identified through the scoping review and national document analysis were subjected to expert validation and contextual refinement. This was achieved through three structured focus group discussions involving 12 subject-matter experts from the Ministry of Health and leading Medical Sciences universities. Participants represented a wide range of professional backgrounds, including health policymakers, academic researchers, emergency and disaster managers, and coordinators specifically engaged in planning for the Arbaeen mass gathering. The diversity of expertise ensured that the checklist would be both evidence-informed and context-sensitive.

The sessions were conducted using an NGT approach, which combined open discussion, individual item rating, and structured voting. At the outset of each session, experts were presented with the preliminary pool of items and were asked to assess each in terms of relevance, clarity, applicability, and cultural appropriateness. Items were discussed one by one, with facilitators ensuring that all participants contributed their perspectives. Following the discussion, participants voted on whether to retain, modify, or remove each item.

During the deliberations, several recurring themes emerged:

- Redundancy and overlap: Many items were found to express similar concepts using different wording. Through consensus, these items were merged or rephrased to enhance clarity.
- Irrelevance or low applicability: Some items were judged too generic, lacking direct relevance to the specific challenges of mass gatherings in the regional context, and were therefore excluded.
- Ambiguity in wording: A smaller set of items (n=28) was removed due to vague phrasing or limited practical applicability.

By the conclusion of the three sessions, 176 items were excluded_148 due to redundancy or irrelevance and 28 due to ambiguity or limited applicability. The final checklist comprised 148 refined items, systematically organized into four domains and eleven subdomains.

The finalized domains and subdomains reflect the multi-dimensional nature of health system preparedness for mass gatherings:

- Health: Communicable and non-communicable diseases, environmental health, and health education.
- Medical Care and Treatment: Disaster management, medical equipment readiness, and emergency response.
- Management and Logistics: Surveillance systems, data and information management, and resource and maintenance management.
- Coordination and Safety: Coordination and cooperation across agencies and safety/security arrangements.

The detailed distribution of items across these domains and subdomains is presented in Table 4. This structure demonstrates the progression from an evidence-based item pool derived from the scoping review to a consensus-driven, contextually adapted checklist, ensuring both scientific rigor and practical utility.

Table 3: Final Domains and Subdomains of the Checklist After Expert Review

Domain	Subdomain	Number of Items
Health	Communicable and Non-Communicable Diseases	14
	Environmental Health	13
	Health Education	9
Medical Care & Treatment	Disaster Management	12
	Medical Care Equipment	11
	Emergencies	10
Management & Logistics	Surveillance Systems	15
	Resource and Maintenance Management	13
	Data and Information Management	14
Coordination & Safety	Coordination and Cooperation	12
	Safety and Security	15
Total		**148 items**

3.3. Phase Three: Quantitative Validation

The third phase aimed to ensure the psychometric soundness of the checklist by systematically evaluating its face validity, content validity, construct validity, and reliability. A multi-step process was conducted as follows:

Face Validity

Face validity was assessed using both qualitative and quantitative approaches. Twelve experts in disaster health management and emergency preparedness independently reviewed the checklist items. For the qualitative assessment, participants provided detailed feedback on item clarity, wording, and applicability in real-world mass gathering contexts. Suggested revisions included simplifying overly technical terms and improving consistency across items.

For the quantitative evaluation, each expert rated item importance on a 5-point Likert scale (1 = not important, 5 = highly important). The impact score was calculated for each item as:

Impact Score = Frequency × Importance

All items achieved an impact score greater than 1.5, confirming that the checklist demonstrated appropriate face validity without the need for item deletion at this stage.

Content Validity

To ensure comprehensiveness and relevance, content validity was examined using CVR and CVI. A large panel of 76 national experts—including policymakers, health system managers, field practitioners, and academics—participated in the evaluation.

- **CVR (necessity):** Experts rated each item as essential, useful but not essential, or not necessary. Following Lawshe's method, items with CVR ≥ 0.56 were retained. All items in the revised checklist met this threshold.

- **CVI (relevance, clarity, and simplicity):** Experts scored each item on a 4-point scale. Items with CVI ≥ 0.79 were retained directly, those between 0.70–0.79 were revised, and items <0.70 were eliminated. The overall mean CVI was 0.98, demonstrating excellent agreement on content clarity and relevance across the expert panel.

Construct Validity

Although the measurement instrument under study can be considered partially formative in nature, most of its domains were designed with reflective indicators that are theoretically expected to covary. Therefore, we applied exploratory factor analysis (EFA) to identify the latent structure of the items and confirmatory factor analysis (CFA) to validate the proposed factor model. The use of EFA and CFA in this context does not imply a purely reflective measurement model, but rather examines whether the hypothesized domains were empirically supported and assesses the degree of internal coherence among items within each domain. This combined approach has also been adopted in several previous validation studies of instruments with mixed formative and reflective features.

The adequacy of the data for factor analysis was confirmed by a Kaiser–Meyer–Olkin (KMO) measure of 0.892 and a statistically significant Bartlett's test of sphericity ($p < 0.001$), indicating the suitability of the data for factor analysis. The EFA using principal component

analysis with Varimax rotation supported the four-domain structure derived from the qualitative phase (Health, Medical Care and Treatment, Management and Logistics, and Coordination and Safety). The scree plot and eigenvalues greater than 1 indicated four factors, consistent with the theoretical framework. All items demonstrated acceptable factor loadings above 0.40, as shown in Table 1. The scree plot is provided in Supplementary Figure 1.

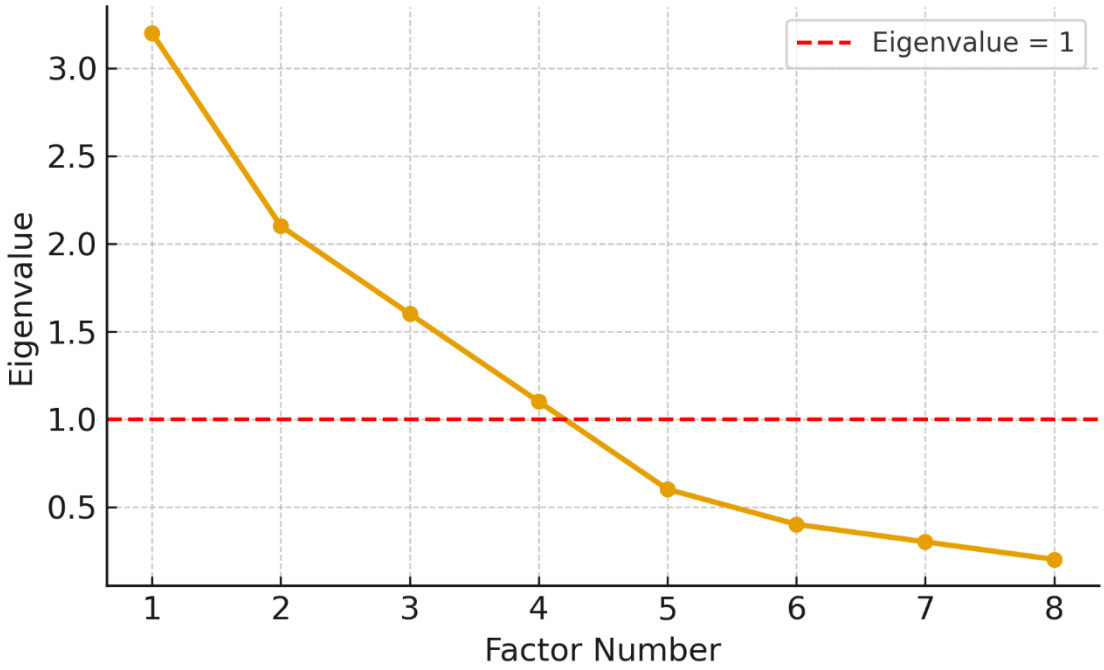

Figure 1: Scree plot of eigenvalues from the exploratory factor analysis (EFA) showing the retention of four factors with eigenvalues greater than 1.

The CFA model further demonstrated acceptable to good fit indices ($\chi^2/df = 2.31$, RMSEA = 0.056, CFI = 0.94, TLI = 0.92). All standardized factor loadings were statistically significant ($p < 0.001$) and above 0.50. These findings confirm the adequacy of the proposed four-factor model and provide strong evidence for the construct validity of the instrument.

Table 4: Factor loadings from exploratory factor analysis (EFA)

Item	Factor 1	Factor 2	Factor 3	Factor 4
Q1	0.68	–	–	–
Q2	0.72	–	–	–
Q3	–	0.65	–	–
Q4	–	0.59	–	–
Q5	–	–	0.74	–
Q6	–	–	0.63	–
Q7	–	–	–	0.81
Q8	–	–	–	0.69

Reliability

Reliability was evaluated in two dimensions:

- Internal consistency: Cronbach's alpha was calculated for the overall checklist and for each domain. The overall Cronbach's alpha was 0.96, indicating excellent internal consistency and suggesting that items within each domain measured related aspects of preparedness.

- Inter-rater reliability: Agreement between evaluators was assessed using the Kappa coefficient, which yielded a value of 0.76, representing substantial agreement according to established thresholds.

Together, these findings confirmed that the checklist demonstrated robust validity and reliability, making it a scientifically sound tool for assessing health system preparedness during mass gatherings (Table 5).

Table 5: Results of Psychometric Evaluation of the Checklist

Evaluation Component	Method/Metric	Result	Interpretation
Face Validity	Impact Score	>1.5 (all items)	Acceptable
Content Validity	Mean CVI (n = 76 experts)	0.98	Excellent content relevance
	Minimum acceptable CVR	≥ 0.56	All retained items met the threshold
Construct Validity	KMO Measure	0.892	Excellent sampling adequacy
	Bartlett's Test (p-value)	<0.001	Factor analysis appropriate
Reliability	Cronbach's Alpha	0.96	High internal consistency
	Kappa Coefficient	0.76	Substantial inter-rater agreement

4. Discussion

This study developed and psychometrically validated the HPAT, with a particular focus on the Arbaeen Pilgrims Walking (APW) event and other large-scale religious gatherings. Using a rigorous mixed-method approach—comprising a scoping review, qualitative expert consultations, and extensive quantitative validation—we finalized a tool with 148 items organized into four major domains (Health, Medical Care and Treatment, Management and Logistics, and Coordination and Safety) and 11 subdomains.

Contribution of the HPAT

The primary contribution of the HPAT lies in its comprehensive coverage of preparedness dimensions, its cultural and contextual relevance, and its psychometric robustness. Previous tools have often focused narrowly on hazard identification or isolated aspects of preparedness,

whereas the HPAT integrates clinical care, public health, coordination mechanisms, logistics, and safety/security measures into a unified framework. By incorporating input from 76 national experts and applying robust quantitative validation methods, the tool provides both theoretical soundness and practical applicability in real-world settings.

Comparison with Existing Tools

When compared with existing instruments, HPAT demonstrates broader coverage and greater methodological rigor:

- SALEM tool (Khan, 2021): Designed primarily for sporting events, SALEM evaluates 17 risk factors but lacks robust psychometric validation and does not comprehensively address critical elements of health system preparedness, such as intersectoral coordination or medical logistics [25]. HPAT, by contrast, underwent extensive psychometric testing (CVI = 0.98; Cronbach's $\alpha = 0.96$) and covers both medical and managerial aspects, ensuring greater reliability and usability.
- WHO–EMRO Jeddah tool (2020): This framework emphasizes infectious hazards and environmental vulnerabilities [8]. While valuable for hazard zoning, it omits domains related to psychological preparedness, resource management, and inter-agency coordination. HPAT addresses these omissions by integrating preparedness measures across both infectious and non-infectious threats, making it more flexible for diverse mass gathering scenarios.
- MGRAT tool (India): Developed qualitatively, the Mass Gathering Risk Assessment Tool includes 22 items across seven domains [26]. While useful for basic risk identification, its scope is substantially narrower than HPAT's 148 items across 11 subdomains, which allows for a more detailed and structured assessment of system readiness and capacity gaps.

Furthermore, our findings are consistent with a recent systematic review that identified over 40 categories of mass gathering health hazards, grouped into five domains [1]. HPAT not only incorporates these domains but expands them by linking hazard identification with preparedness strategies, policy frameworks, and operational readiness, thereby bridging the gap between risk recognition and actionable response planning.

Unique Features and Advantages

HPAT demonstrates several advantages over existing tools:

1. Multidimensional Structure – Unlike risk-only tools, HPAT balances clinical, public health, logistical, and managerial domains.
2. Psychometric Validation – The inclusion of face validity, content validity, construct validity, and reliability testing enhances the scientific rigor of the tool, a feature absent in most prior frameworks.

3. Contextual Adaptability – Designed with input from local and national experts, HPAT addresses religious and cultural dimensions of gatherings in the Middle East, where traditional tools often fail to capture unique contextual challenges.
4. Policy and Practice Integration – The tool provides action-oriented outputs that support evidence-based planning, resource allocation, and training programs.

Implications for Policy and Practice

The development and validation of the HPAT carry important implications for both policy and practice. By providing a standardized and psychometrically robust framework, the tool enables health authorities to systematically identify strengths and weaknesses within the health system. Such structured assessment facilitates the more efficient allocation of resources and the prioritization of capacity-building initiatives, ensuring that areas with the greatest vulnerabilities receive focused attention. Moreover, the HPAT emphasizes the necessity of intersectoral and interagency collaboration, an aspect often overlooked in existing tools but critical for the complex coordination required during large-scale mass gatherings. Its comprehensive approach also supports preparedness for a wide spectrum of threats, from communicable disease outbreaks to non-communicable emergencies and logistical failures. In practice, the tool can guide evidence-based planning, improve training programs, and strengthen real-time decision-making processes, ultimately contributing to safer and more resilient health systems during events such as the Arbaeen Pilgrimage.

5. Conclusion

This study developed and validated the HPAT as a comprehensive and context-specific instrument for assessing preparedness in mass gatherings such as the Arbaeen Pilgrimage. By integrating evidence from a scoping review, expert consultations, and quantitative validation, the tool was refined to include148 items across four domains and eleven subdomains. Unlike previous tools, the HPAT offers a multidimensional and psychometrically sound framework that addresses both operational and cultural needs, thereby filling an important gap in preparedness assessment. Future research should examine its applicability in diverse settings and further explore its predictive validity in real-world mass gatherings.

6. Declaration:

Authors' Contribution:

Fariba Ahmadi: Performed data gathering; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools, or data; Wrote the paper.
Zohreh Ghomian, Azadeh Tahernejad, and **Sakineh Rakhshanderou:** Conceived and designed the study and experiments; Contributed reagents, materials, analysis tools, or data; Supervised and verified the findings; Finalized the manuscript.

Conflict of Interest:

The authors declare that they have no competing interests.

Funding:

The authors received no financial support for the research.

Ethics Approval and Consent to Participate:

This study received the required ethics approval from the Shahid Beheshti University of Medical Sciences Research Ethics Committee, Tehran, Iran, with ethical code no. IR.SBMU.PHNS.REC.1401.102. All study participants provided written informed consent to participate. During focus group discussions (FGDs), written consent was obtained from participants before recording. Participants were informed that their participation was voluntary, and confidentiality was assured. All methods were carried out in accordance with relevant guidelines and regulations.

Availability of Data and Materials:

All data generated or analyzed during this study are included in this published article.

7. References

1. Tavan, A., et al., *Risks threatening the health of people participating in mass gatherings: A systematic review.* 2019. **8**.
2. Lee, W.V., *Historical global analysis of occurrences and human casualties of extreme temperature events (ETEs).* Natural hazards, 2014. **70**: p. 1453-1505.
3. de Paiva Teixeira, L.E.P., et al., *Psychological impacts related to stress and fear during the COVID-19 pandemic: cardiovascular diseases, diabetes and psychological disorders as risk factors.* World Journal of Neuroscience, 2020. **10**(4): p. 191-205.
4. Almehmadi, M. and J.S. Alqahtani. *Healthcare research in mass religious gatherings and emergency management: a comprehensive narrative review*. in *Healthcare*. 2023. MDPI.
5. Zohreh Ghomian, Z.E., Arezoo Dehghani, *Human Mass Gatherings and Communicable Diseases: COVID-19.* Journal of Culture and Health Promotion of the Academy of Medical Sciences, 2023. **4**(1).
6. Farahani, A.J. and S.R.H. Zijoud, *Research and researchers of hygiene, medicine and health field in the service of Arbaeen walk.* Journal of Military Medicine, 2019. **21**(6): p. 547-548.
7. Safarpour, H., et al., *Prevalence of Influenza Among Hajj Pilgrims: A Systematic Review and Meta-Analysis.* Disaster medicine and public health preparedness, 2022. **16**(3): p. 1221-1228.
8. Bieh, K.L., et al., *Guidance for health risk assessment at recurrent mass gatherings: The Jeddah tool framework.* 2021. **36**(3): p. 348-353.
9. Karami, M., et al., *Public health threats in mass gatherings: A systematic review.* Disaster medicine and public health preparedness, 2019. **13**(5-6): p. 1035-1046.
10. Bieh, K., et al., *Health risk assessment at mass gatherings: a report of the camel festival in Saudi Arabia.* 2019. **25**(9).
11. World Health Organization, *Key planning recommendations for mass gatherings in the context of COVID-19: interim guidance, first issued 14 February 2020, third update 4 November 2021*. 2021, World Health Organization.
12. Khankeh, H.R., et al., *Effects hospital incident command system establishment on disaster preparedness of Tehran hospitals affiliated to law enforcement staff under simulated conditions.* Health in emergencies and disasters quarterly, 2018. **3**(4): p. 207-214.
13. Arab M, Z.H., Akbari Haghighi F, Ravangard R, *A study on the executive managers' knowledge and performance, and their hospitals preparedness against earthquake events and their relationships at public hospitals (affiliated by Tehran University of Medical Sciences (TUMS) 2005-2006.* Journal of Health Administration, 2009. **11**(34): p. 7-14.
14. World Health Organization, *Public health for mass gatherings: key considerations*. 2015, World Health Organization.
15. Crano, W.D., Brewer, M. B., & Lac, A,, *Principles and methods of social research (3rd ed.)*. 2015, London and New York, NY: Routledge. .
16. Delany, C., J. Crilly, and J. Ranse, *Drug and alcohol related patient presentations to emergency departments during sporting mass-gathering events: an integrative review.* Prehospital and disaster medicine, 2020. **35**(3): p. 298-304.

17. McCann, C.M., et al., *Resilience in the health professions: A review of recent literature.* International Journal of Wellbeing, 2013. **3**(1).
18. MirKazemi, S.O., S. Rastgoo, and M. Khorashadizadeh, *Designing an Instrument to Assess the Factors Affecting Human Resource Productivity in Iran Sports Federations.* Sport Management and Development, 2017. **6**(2): p. 1-15.
19. Connell, J., et al., *The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL).* Quality of life research, 2018. **27**: p. 1893-1902.
20. Dehghani, A., Z. Ghomian, and Z. Eskandari, *Design and psychometrics of risk assessment tool for makeshift hospitals: Focusing on pandemics.* Heliyon, 2023. **9**(4).
21. Larson, R.J.J.o.g., *Thirty years of research on the subjective well-being of older Americans.* 1978. **33**(1): p. 109-125.
22. Tahernejad, A., et al., *Investigating the conceptual model of the formation of psychological resilience in farmers affected by droughts in Iran using structural equation modeling.* BMC Public Health, 2025. **25**(1): p. 1883.
23. Thomas, E. and J.K. Magilvy, *Qualitative rigor or research validity in qualitative research.* Journal for specialists in pediatric nursing, 2011. **16**(2).
24. Russo, J., et al., *Longitudinal assessment of quality of life in acute psychiatric inpatients: reliability and validity.* 1997. **185**(3): p. 166-175.
25. Khan, A.A., et al., *Mass gathering medicine in soccer leagues: a review and creation of the SALEM Tool.* 2021. **18**(19): p. 9973.
26. Sharma, U., B. Desikachari, and S. Sarma, *Content validity of the newly developed risk assessment tool for religious mass gathering events in an Indian setting (Mass Gathering Risk Assessment Tool-MGRAT).* Journal of family medicine and primary care, 2019. **8**(7): p. 2207-2211.