Received: 22 March 2025 / Accepted: 15 November 2025 / Published online: 30 December 2025

DOI 10.34689/SH.2025.27.6.011

UDC 616.6-08-055

METHODOLOGICAL STUDY VALIDATION OF THE INSTRUMENT «ORGANIZATIONAL ASPECTS OF URO-ANDROLOGICAL CARE» FOR USE BY HEALTHCARE ORGANIZATION SPECIALISTS

Bayan B. Orazayeva[1], https://orcid.org/0009-0005-0374-7884

Zaituna A. Khismetova[1], https://orcid.org/0000-0001-5937-3045

Merkhat N. Akkaliyev[1], https://orcid.org/0000-0003-3122-7411

[1] NCJSC «Semey Medical University», Semey, Republic of Kazakhstan.

Abstract

Introduction: Sociological questionnaires play a crucial role in most studies in the field of healthcare, as the validity, reproducibility, and interpretability of research findings largely depend on the quality of the measurement instrument. Incorrectly developed and/or non-validated instruments may lead to systematic measurement errors, distortion of the true characteristics of the phenomenon under study, and, consequently, to inappropriate organizational and managerial decisions. The lack of internationally validated questionnaires aimed at identifying organizational barriers in the provision of uro-andrological care, as well as at building an evidence base for optimizing healthcare services for men, highlights the need for the development and validation of such an instrument.

Aim of the Study: to scientifically substantiate the validation process of a multi-aspect questionnaire entitled «Organizational Aspects of Uro-Andrological Care» for its subsequent use by healthcare organization specialists.

Materials and Methods: A multi-stage validation procedure was conducted, including the following phases: conceptual (theoretical) validity; face validity; content validity; cognitive validity (pilot study); assessment of internal consistency reliability; and finalization of the questionnaire.

Results: The results of the conceptual and face validity assessments demonstrated favorable outcomes. High content validity was confirmed by a panel of eight experts (I-CVI = 0.875–1.00; S-CVI/Ave = 0.987–1.00; S-CVI/UA = 0.895–1.00). The pilot study (n = 10) indicated a high level of interpretative clarity of the questionnaire items. The internal consistency of the instrument was satisfactory (Cronbach's alpha = 0.739 for 23 included items), indicating an acceptable level of reliability at the pilot testing stage.

Conclusions: The developed multi-aspect questionnaire «Organizational Aspects of Uro-Andrological Care» underwent a comprehensive, multi-level validation process. These findings confirm the methodological soundness and scientific applicability of the instrument for use by healthcare organizers in further research aimed at examining organizational aspects of male reproductive health, particularly at the regional level.

Keywords: *multi-aspect questionnaire; instrument; validation; uro-andrological care; reproductive health; men; region.*

For citation:

Orazayeva B.B., Khismetova Z.A., Akkaliyev M.N. Methodological study. Validation of the instrument «organizational aspects of uro-andrological care» for use by healthcare organization specialists // *Nauka i Zdravookhranenie* [Science & Healthcare]. 2025. Vol.27 (6), C.91-100. *doi 10.34689/SH.2025.27.6.011*

Резюме

МЕТОДОЛОГИЧЕСКАЯ ВАЛИДАЦИЯ ИНСТРУМЕНТА «ОРГАНИЗАЦИОННЫЕ АСПЕКТЫ УРОАНДРОЛОГИЧЕСКОЙ МЕДИЦИНСКОЙ ПОМОЩИ» ДЛЯ ИСПОЛЬЗОВАНИЯ СПЕЦИАЛИСТАМИ МЕДИЦИНСКИХ ОРГАНИЗАЦИЙ

Баян Б. Оразаева[1], https://orcid.org/0009-0005-0374-7884

Зайтуна А. Хисметова[1], https://orcid.org/0000-0001-5937-3045

Мерхат Н. Аккалиев[1], https://orcid.org/0000-0003-3122-7411

[1] НАО «Медицинский Университет Семей», Семей, Республика Казахстан.

Введение: Социологические анкеты играют решающую роль в большинстве исследований в области здравоохранения, поскольку достоверность, воспроизводимость и интерпретируемость результатов исследований во многом зависят от качества измерительного инструмента. Неправильно разработанные и/или невалидированные инструменты могут привести к систематическим ошибкам измерения, искажению истинных характеристик изучаемого явления и, следовательно, к неадекватным организационным и управленческим решениям. Отсутствие международно валидированных анкет, направленных на выявление организационных барьеров в предоставлении

уроандрологической помощи, а также на создание доказательной базы для оптимизации медицинских услуг для мужчин, подчеркивает необходимость разработки и валидации такого инструмента.

Цель исследования: научно обосновать процесс валидации многоаспектной анкеты под названием «Организационные аспекты уроандрологической помощи» для ее последующего использования специалистами по организации здравоохранения.

Материалы и методы: Была проведена многоэтапная процедура валидации, включающая следующие этапы: концептуальная (теоретическая) валидность; внешняя валидность; содержательная валидность; когнитивная валидность (пилотное исследование); оценка надежности внутренней согласованности; и окончательная доработка анкеты.

Результаты: Результаты оценки концептуальной и внешней валидности продемонстрировали благоприятные результаты. Высокая содержательная валидность была подтверждена группой из восьми экспертов (I-CVI = 0,875–1,00; S-CVI/Ave = 0,987–1,00; S-CVI/UA = 0,895–1,00). Пилотное исследование (n = 10) показало высокий уровень интерпретационной ясности пунктов анкеты. Внутренняя согласованность инструмента была удовлетворительной (альфа Кронбаха = 0,739 для 23 включенных пунктов), что указывает на приемлемый уровень надежности на этапе пилотного тестирования.

Выводы: Разработанная многоаспектная анкета «Организационные аспекты уроандрологической помощи» прошла всесторонний многоуровневый процесс валидации. Полученные результаты подтверждают методологическую обоснованность и научную применимость инструмента для использования организаторами здравоохранения в дальнейших исследованиях, направленных на изучение организационных аспектов репродуктивного здоровья мужчин, особенно на региональном уровне.

Ключевые слова: *многоаспектный опросник; инструмент; валидация; уроандрологическая помощь; репродуктивное здоровье; мужчины; регион.*

Для цитирования:

Оразаева Б.Б., Хисметова З.А., Аккалиев М.Н. Методологическая валидация инструмента «Организационные аспекты уроандрологической медицинской помощи» для использования специалистами медицинских организаций // Наука и Здравоохранение. 2025. Vol.27 (6), C.91-100. *doi 10.34689/SH.2025.27.6.011*

Түйіндеме

МЕТОДОЛОГИЧЕСКАЯ ВАЛИДАЦИЯ ИНСТРУМЕНТА «ОРГАНИЗАЦИОННЫЕ АСПЕКТЫ УРОАНДРОЛОГИЧЕСКОЙ МЕДИЦИНСКОЙ ПОМОЩИ» ДЛЯ ИСПОЛЬЗОВАНИЯ СПЕЦИАЛИСТАМИ МЕДИЦИНСКИХ ОРГАНИЗАЦИЙ

Баян Б. Оразаева[1], https://orcid.org/0009-0005-0374-7884

Зайтуна А. Хисметова[1], https://orcid.org/0000-0001-5937-3045

Мерхат Н. Аккалиев[1], https://orcid.org/0000-0003-3122-7411

[1] «Семей медицина университеті» КеАҚ, Семей, Қазақстан Республикасы.

Введение: Социологические анкеты играют решающую роль в большинстве исследований в области здравоохранения, поскольку достоверность, воспроизводимость и интерпретируемость результатов исследований во многом зависят от качества измерительного инструмента. Неправильно разработанные и/или невалидированные инструменты могут привести к систематическим ошибкам измерения, искажению истинных характеристик изучаемого явления и, следовательно, к неадекватным организационным и управленческим решениям. Отсутствие международно валидированных анкет, направленных на выявление организационных барьеров в предоставлении уроандрологической помощи, а также на создание доказательной базы для оптимизации медицинских услуг для мужчин, подчеркивает необходимость разработки и валидации такого инструмента.

Цель исследования: научно обосновать процесс валидации многоаспектной анкеты под названием «Организационные аспекты уроандрологической помощи» для ее последующего использования специалистами по организации здравоохранения.

Материалы и методы: Была проведена многоэтапная процедура валидации, включающая следующие этапы: концептуальная (теоретическая) валидность; внешняя валидность; содержательная валидность; когнитивная валидность (пилотное исследование); оценка надежности внутренней согласованности; и окончательная доработка анкеты.

Результаты: Результаты оценки концептуальной и внешней валидности продемонстрировали благоприятные результаты. Высокая содержательная валидность была подтверждена группой из восьми экспертов (I-CVI = 0,875–1,00; S-CVI/Ave = 0,987–1,00; S-CVI/UA = 0,895–1,00). Пилотное исследование (n = 10) показало высокий уровень интерпретационной ясности пунктов анкеты. Внутренняя согласованность инструмента была удовлетворительной (альфа Кронбаха = 0,739 для 23 включенных пунктов), что указывает на приемлемый уровень надежности на этапе пилотного тестирования.

Выводы: Разработанная многоаспектная анкета «Организационные аспекты уроандрологической помощи» прошла всесторонний многоуровневый процесс валидации. Полученные результаты подтверждают

методологическую обоснованность и научную применимость инструмента для использования организаторами здравоохранения в дальнейших исследованиях, направленных на изучение организационных аспектов репродуктивного здоровья мужчин, особенно на региональном уровне.

Ключевые слова: *многоаспектный опросник; инструмент; валидация; уроандрологическая помощь; репродуктивное здоровье; мужчины; регион.*

Дәйексөз үшін:

Оразаева Б.Б., Хисметова З.А., Аккалиев М.Н. Методологическая валидация инструмента «Организационные аспекты уроандрологической медицинской помощи» для использования специалистами медицинских организаций // Ғылым және Денсаулық сақтау. 2025. Vol.27 (6), С.91-100. *doi 10.34689/SH.2025.27.6.011*

Relevance

At present, a steady upward trend in the prevalence of infertile marriages has been observed, representing a serious medico-social problem [19]. Despite the fact that 45.0–52.0% of infertility cases are associated with disorders of the male reproductive system [6] [14], issues related to male reproductive health and the state of uro-andrological care have long remained outside the focus of professional attention. A number of studies emphasize the existence of a crisis [8] and the so-called «relative neglect» [25] with regard to male fertility health, highlighting the need for organizational and functional modernization of uro-andrological services [1], [9], [18].

In this context, there is an urgent need to study the opinions and attitudes of urologists and andrologists toward the organizational aspects of uro-andrological care as an integrated system. However, to date, no unified, standardized, and validated instrument exists that adequately meets this requirement.

Numerous studies have employed internationally recognized questionnaires; however, these instruments are characterized by a narrow thematic focus, addressing specific conditions such as varicocele [26], idiopathic infertility [5], [16], erectile dysfunction [13], assessment of quality of life among men experiencing fertility difficulties [3], and evaluation of primary healthcare physicians' knowledge regarding the management of urological diseases [22], among others. None of these questionnaires is designed to provide a comprehensive assessment of the organizational aspects of uro-andrological care.

Moreover, one of the key objectives of sociological research is to establish the validity and reliability of the measurement instrument. Validity reflects the extent to which a measurement tool assesses what it is intended to measure, while reliability indicates the degree of reproducibility of the results obtained through measurement [4], [17]. In light of the above, there is an objective need for the development and validation of an instrument capable of adequately assessing the relevant aspects of the problem under investigation.

Aim of the Study: to scientifically substantiate the validation process of a multi-aspect questionnaire entitled «Organizational Aspects of Uro-Andrological Care» for its subsequent use by healthcare organization specialists.

Materials and Methods: To ensure methodological rigor and improve the quality of the multi-aspect questionnaire designed to assess the organizational aspects of uro-andrological care at the regional level, a multi-stage validation procedure was conducted. This process included the following stages:

1. conceptual (theoretical) validity;
2. face validity;
3. content validity;
4. cognitive validity (pilot study);
5. assessment of internal consistency reliability;
6. finalization of the questionnaire.

The preliminary version of the instrument, i.e., the prototype questionnaire entitled «Organizational Aspects of Uro-Andrological Care», was developed in two languages – Russian and Kazakh – and consisted of 45 items. For the professional translation of the instrument into the Kazakh language, two certified linguists who were native speakers and had practical experience working with medical terminology were involved.

Following the assessment of conceptual validity, which included a detailed analysis of the questionnaire's content structure – namely, the removal of duplicate items, exclusion of questions with overlapping meanings, and elimination of formulations that were not aligned with the study objective – the final version of the questionnaire was established on May 13, 2025. The finalized version comprised 38 items.

At the initial stage, between May 14 and May 16, 2025, an assessment of face validity was conducted to evaluate the clarity, comprehensibility, and unambiguity of the questionnaire items. Representatives of the target population – practicing specialists, namely urologists-andrologists (n=5) – participated in this procedure. Participants were asked to assess each item according to criteria of clarity, logical consistency, and relevance to the stated topic. No substantial comments were identified as a result of the face validity assessment.

The expert evaluation of the instrument was carried out in accordance with the methodologies proposed by Waltz and Bausell (1981) and Lynn (1986) during the period from May 25 to June 3, 2025. Content validity was assessed by eight independent experts using a four-point rating scale. Based on the expert assessment, the following indices were calculated:

I-CVI (Item-Level Content Validity Index) – the proportion of experts who rated an item as 3 or 4; a threshold value of I-CVI≥0.78 is considered acceptable when the number of experts is ≥6–10;

S-CVI/Ave (Scale Content Validity Index, Average) – the average I-CVI across all items of the questionnaire; a threshold value of S-CVI/Ave≥0.90 is considered acceptable for ≥6–10 experts;

S-CVI/UA (Scale Content Validity Index, Universal Agreement) – the proportion of items for which all experts

were in complete agreement; a threshold value of S-CVI/UA ≥0.80 is considered acceptable for ≥6–10 experts.

According to the analysis, all three indices reached the required levels of content validity. Based on expert comments, selected item formulations and/or response options were refined.

Following these revisions, a pilot study was conducted between June 9 and June 13, 2025, among a sample of urologists-andrologists (n=10) representing the target population (pilot testing on a small sample). The primary aim of the pilot testing was to assess the applicability and practical feasibility of the instrument. The average time required to complete the questionnaire ranged from 15 to 20 minutes. No structural errors were identified, with the exception of an inaccuracy in item No.34, which confirmed the overall suitability of the instrument for the main study. Item No.34 and its response options were subsequently rephrased. Given that the sample size was ≤50, conducting factor analysis was deemed inappropriate.

To assess the internal consistency, or reliability, of the questionnaire scales, Cronbach's alpha coefficient was calculated. The analysis was performed using IBM SPSS Statistics version 20, both for the questionnaire as a whole and for individual items. An alpha value of ≥0.70 is considered acceptable for social and medical research. The Cronbach's alpha value calculated based on the pilot sample should be interpreted as preliminary.

Based on the results of the above-described validation stages, the final version of the multi-aspect questionnaire «Organizational Aspects of Uro-Andrological Care» was established. In addition, a certificate of registration in the State Register of Rights to Copyright-Protected Objects was obtained (No. 62500, dated September 29, 2025).

Results: Questionnaire-based studies are neither purely quantitative nor purely qualitative in nature; rather, they are designed to collect data of various types depending on the research questions being addressed.

For the purpose of developing and validating the multi-aspect questionnaire «Organizational Aspects of Uro-Andrological Care» intended for specialists involved in organizational issues of male reproductive health, the following procedures were carried out.

Conceptual (Theoretical) Validity of the Questionnaire

At this stage of validation, content optimization of the preliminary version of the questionnaire was performed. This process included a detailed analysis of item wording, removal of overlapping and duplicate questions, and exclusion of non-relevant items that did not correspond to the study objectives. As a result, the number of items was reduced from 45 to 38.

The multi-aspect questionnaire «Organizational Aspects of Uro-Andrological Care», designed for specialists involved in male reproductive health issues, was developed for subsequent use in the identification, scientific justification, and prioritization of key directions for the development of the male reproductive health protection system at the regional level.

The questionnaire consists of two main sections and several sub-sections:

1.Sociodemographic and professional characteristics;

2.Organization of uro-andrological care, including the following domains:

– organization of the uro-andrological care system for the regional population (patient routing, interdisciplinary coordination);

– accessibility of urological and andrological care for the regional population;

– rehabilitation and dispensary follow-up of men with urological and andrological conditions;

– material and technical resources for the diagnosis and treatment of male reproductive health disorders;

– monitoring of men's health status and collection of patient feedback;

– workforce capacity and professional workload of medical specialists;

– organizational aspects of diagnostics and screening of male reproductive health at the regional level;

– reproductive behavior and gender attitudes toward childbearing;

– informatization and digitalization of the uro-andrological care system;

– scientific, educational, and preventive programs for healthcare professionals and the regional population.

The above-mentioned domains of the questionnaire enable comprehensive identification of key organizational-processual and sociocultural barriers affecting the functioning of the existing uro-andrological care system in the region, as well as assessment of their impact on the accessibility, quality, and effectiveness of specialized medical care.

Face Validity

Face validity represents one of the important stages of instrument validation and is based on the external appearance, format, and structural organization of the questionnaire. This stage is aimed at assessing the clarity, correctness, and unambiguity of item wording, as well as the overall comprehensibility of the questions for the target audience.

Taking into account feedback not only from experts but also from members of the target research population is considered essential [7], as the evaluation provided by potential respondents contributes to improving the acceptability, relevance, and overall quality of the instrument [32].

The face validity assessment involved practicing specialists – urologists-andrologists (n=5) – representing the target population but not included in the expert panel. Participants were selected using a random sampling method.

The respondents were asked to evaluate each questionnaire item according to the following criteria:

1) clarity (yes/no);
2) absence of ambiguity (yes/no);
3) appropriateness of wording (yes/no).

Based on the results of the face validity assessment, no substantial comments or concerns were identified. The evaluation was documented in written form using a standardized assessment sheet.

Content Validity (Expert Evaluation)

There is no single standardized or universally approved format for expert evaluation, as each instrument and

questionnaire is inherently unique. Research objectives, questionnaire structure, and item content vary; therefore, the expert evaluation form should accurately reflect the specific content of the instrument being assessed. In the present study, the ultimate purpose of the questionnaire «Organizational Aspects of Uro-Andrological Care» was to develop and scientifically substantiate priority directions for the male reproductive health protection system.

In addition, internationally recognized methodological frameworks – such as those proposed by Waltz and Bausell (1981) [31], Lynn (1986) [20], and Polit and Beck (2006) [24] – provide general principles for conducting expert evaluation rather than a fixed template.

The aim of the expert evaluation was to determine the content validity of the questionnaire, that is, the extent to which each item corresponds to the study objectives, is clear and appropriate for the target population, and is supported by high-quality response options. Unlike other forms of validity, content validity plays a decisive role in the interpretation of research findings and their practical application [20].

Accordingly, experts selected for content validation were required to possess in-depth knowledge of the subject matter and prior experience in the validation of sociological instruments. Lynn (1986) recommended involving between 3 and 10 experts in the content validation process [20].

For the assessment of content validity of the multi-aspect questionnaire, an expert panel consisting of eight specialists with academic degrees was formed. The panel included practicing urologists as well as scientific and methodological experts from different regions of the country, all of whom had professional experience in both academic research and practical healthcare. This composition ensured a balance between knowledge of regional specificities in uro-andrological care delivery and an independent scientific and methodological perspective.

Each expert was provided with the questionnaire, a description of the study objectives, an expert evaluation form, and detailed instructions for its completion.

Experts evaluated each questionnaire item using a four-point Likert scale (from 1 to 4) according to the following assessment domains:

Objective relevance – the extent to which an individual item reflects the stated objective of the questionnaire and measures the intended construct:

1 – not relevant, does not reflect the questionnaire objective, requires complete revision;

2 – partially relevant, requires substantial revision;

3 – generally relevant, minor revisions possible;

4 – fully relevant, no revision required.

Clarity – the degree to which an item is clearly, grammatically, and logically formulated and understandable to potential respondents:

1 – unclear or ambiguous, requires complete revision;

2 – partially clear, requires substantial revision;

3 – generally clear, minor revisions possible;

4 – completely clear, no revision required.

Appropriateness – the extent to which an item is appropriate within the professional, cultural, and contextual framework of the target population:

1 – inappropriate, requires complete revision;

2 – partially appropriate, requires substantial revision;

3 – generally appropriate, minor revisions possible;

4 – fully appropriate, no revision required.

Quality of response options – the accuracy, completeness, and logical structure of the proposed response options, enabling respondents to adequately express their opinions or select appropriate answers:

1 – incorrect and/or incomplete, requires complete revision;

2 – partially correct, requires substantial revision;

3 – generally correct, minor revisions possible;

4 – fully correct, no revision required.

Experts were also given the opportunity to provide qualitative comments and recommendations aimed at optimizing item wording and improving the overall quality of the questionnaire.

In accordance with the methodology proposed by Waltz and Bausell (1981), expert ratings were dichotomized: items rated as 3 or 4 were assigned a value of «1», indicating relevance, whereas items rated as 1 or 2 were assigned a value of «0», indicating non-relevance and the need for further revision.

Based on the expert ratings and subsequent dichotomization, the Item-Level Content Validity Index (I-CVI) was calculated for each questionnaire item. The I-CVI represents the proportion of experts who considered a given item to be relevant and was calculated using the following formula:

I-CVI = (number of experts rating the item as 3 or 4) / (total number of experts)

According to methodological recommendations, an I-CVI value ≥0.78 is considered acceptable when the number of experts ranges from 6 to 10; items with I-CVI <0.78 require revision [27].

In the present study, the vast majority of questionnaire items demonstrated an I-CVI of 1.00, indicating a high level of relevance of both item wording and response options. For items No. 18, 19, 27, and 28, the I-CVI for objective relevance was 0.875, which exceeds the recommended threshold of ≥0.78 and indicates acceptable content validity (Figure 1. I-CVI Objective Relevance). No items with I-CVI values below 0.78 were identified; therefore, no further item revision was required.

Next, the Scale-Level Content Validity Index (S-CVI) was calculated to assess the overall content validity of the questionnaire. It is important to note that two types of S-CVI are distinguished in the methodological literature [11]:

1) S-CVI/Ave (Scale Content Validity Index, Average) – reflects the average content validity of the entire scale and is calculated as the mean value of all item-level content validity indices (I-CVI). Formula for S-CVI/Ave:

S-CVI/Ave = (sum of I-CVI values) / (total number of items)

2) S-CVI/UA (Scale Content Validity Index, Universal Agreement) – represents the proportion of items for which universal agreement among experts was achieved, that is, the share of items rated as 3 or 4 by all experts. Formula for S-CVI/UA:

S-CVI/UA = (number of items rated 3 or 4 by all experts) / (total number of items)

The recommended threshold values for acceptable content validity are S-CVI/Ave >0.90 and S-CVI/UA >0.80.

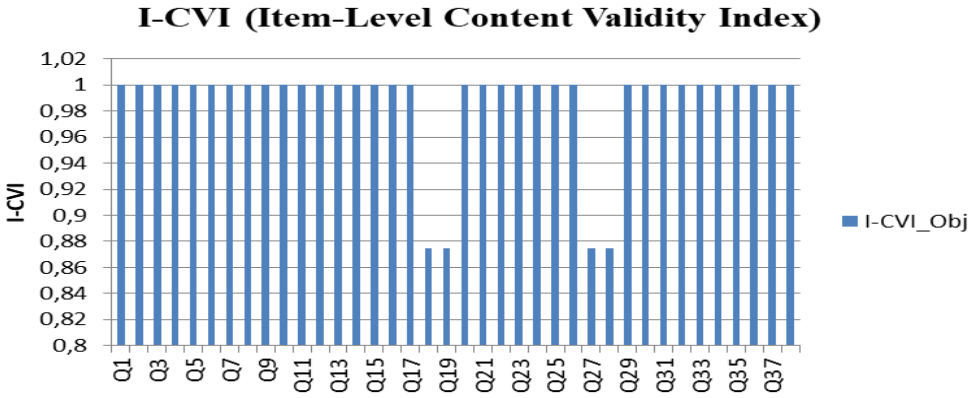

Figure 1. I-CVI Objective Relevance

The results of the calculations indicate a high level of expert agreement across all questionnaire domains. Specifically, for the domains «Clarity», «Appropriateness», and «Quality of Response Options», both indices reached a value of 1.00. For the domain «Objective Relevance», the values of S-CVI/Ave and S-CVI/UA were 0.987 and 0.895, respectively. Thus, all domains of the instrument demonstrated excellent content validity, and the questionnaire structure can be considered methodologically and conceptually sound.

During the content validation process, experts conducted not only a quantitative assessment but also a qualitative evaluation of each questionnaire item, providing detailed comments and recommendations. These comments addressed aspects such as clarity of wording, terminological accuracy, semantic refinement of items, and clarification of response options. All expert feedback was systematically organized and subjected to qualitative analysis. Based on these recommendations, selected items and/or response options were rephrased, which contributed to improved content validity and enhanced cognitive accessibility of the instrument.

Cognitive Validity (Pilot Study)

A pilot study is a preliminary investigation conducted prior to the actual survey or experiment and is designed to examine various aspects of the research process. In sociological research, pilot studies are typically carried out to assess the feasibility and suitability of a research instrument – namely, a questionnaire – before its application in a larger-scale study. In addition, pilot testing allows for the evaluation of key parameters of the chosen study design and methodology, participant selection criteria, and other operational strategies [21].

Conducting a large-scale survey is generally resource-intensive. Any substantial errors identified during a full-scale survey are highly likely to result in inefficient use of resources, including time, workforce, and financial costs. Therefore, a pilot study serves as an essential preparatory stage prior to the main survey, ensuring that the primary study can be implemented with realistic prospects of achieving its research objectives.

The present instrument is intended for use in the implementation of the third objective of the scientific start-up project «Model of a Digital Platform for Improving the Male Reproductive Health Protection System», namely, to identify current barriers within the existing healthcare organization related to male reproductive health. To achieve this objective, a questionnaire survey is planned among urologists and andrologists in the eastern region of the Republic of Kazakhstan (hereinafter referred to as the RK).

One of the fundamental considerations in planning a pilot study is the determination of the minimum sample size required to assess instrument reliability. According to official data from the Republican State Enterprise on the Right of Economic Management «Salidat Kairbekova National Scientific Center for Healthcare Development» of the Ministry of Health of the Republic of Kazakhstan, as of 2024, the number of practicing urologists and andrologists was 12 in the Abai Region, 18 in the East Kazakhstan Region and 23 in the Pavlodar Region.

Given that the target population in the present study was relatively small (n=53), it was decided to recruit 10 specialists to participate in the pilot testing of the questionnaire. Respondents were selected using a random sampling method, which ensured sample representativeness.

At this stage, the assessment focused on the clarity of item wording, verification of the logical sequence of questions, identification of potential difficulties in questionnaire completion, and determination of the time required to complete the questionnaire.

The allotted time for comprehensive and accurate questionnaire completion (15–20 minutes) was deemed sufficient. Overall, respondents did not report difficulties in completing the questionnaire. However, 7 out of 10 respondents (70%) experienced difficulty when answering Item No. 34: «How do you assess the level of organization of preventive examinations and screening programs for men compared with similar programs for women?»

The original response options were as follows:

1.significantly higher

2.somewhat higher

3.at the same level

4.somewhat lower

5.significantly lower

Considering the difficulties identified in interpreting the original wording of Item No. 34, and in order to enhance clarity and improve the validity of responses, a decision was made to modify the original formulation as follows: «In your opinion, the level of organization of preventive examinations and screening for male reproductive health is:»

The revised response options were:

1.significantly lower than that for women

2.somewhat lower than that for women

3.approximately at the same level

4.higher than that for women

5.difficult to answer

Thus, based on the results of the pilot study conducted on a small sample, no substantial content-related concerns were identified by respondents. The overall structure of the instrument, as well as the wording of questionnaire items and corresponding response options, was approved for subsequent use.

Assessment of Internal Consistency Reliability (Cronbach's Alpha)

The internal consistency of the questionnaire was assessed using Cronbach's alpha coefficient, which describes the extent to which all items of a test measure the same underlying construct and, consequently, reflects the interrelatedness of the items within the instrument. Internal consistency should be established prior to the use of a test for research purposes in order to ensure its validity. In addition, reliability estimates indicate the magnitude of measurement error in a test. In other words, reliability may be interpreted as the correlation of the test with itself. Cronbach's alpha coefficient ranges from 0 to 1 [30], with values closer to 1 indicating higher reliability of the scale.

The calculation of Cronbach's alpha on the pilot sample (n=10) was performed to identify potentially problematic questionnaire items and should be regarded solely as a preliminary indicator of the internal consistency of the instrument.

Cronbach's alpha was calculated and analyzed for the questionnaire as a whole as well as for individual items. The literature reports various acceptable alpha values, typically ranging from 0.70 to 0.95 [30]. Low alpha values may be associated with a small number of items, weak inter-item correlations, or heterogeneity of the measured constructs.

In accordance with the recommendations of Jum C. Nunnally (1978) [23] and Keith S. Taber (2018) [29], the following classification and interpretation of Cronbach's alpha values were applied:

α ≥ 0.90 – excellent internal consistency;

α = 0.80–0.89 – good internal consistency;

α = 0.70–0.79 – satisfactory internal consistency;

α < 0.70 – revision required.

In line with psychometric requirements, only scale (ordinal) items were included in the calculation of Cronbach's alpha. Sociodemographic («passport») questions, as well as items that were non-measurement in nature and not suitable for internal consistency assessment, were excluded from the analysis.

The obtained Cronbach's alpha value was 0.739 (based on 23 included items), which is considered a satisfactory level of internal consistency for social and medical research and indicates an acceptable level of instrument reliability at the pilot testing stage.

Cronbach's alpha values calculated after deleting individual items ranged from 0.691 to 0.818 (Table 1. Item–Total Statistics). Higher alpha values observed after item deletion indicate relatively weaker consistency of specific items with the overall scale.

Table 1.

Item–Total Statistics

Item number	Scale mean if item deleted	Scale variance if item deleted	Corrected item–total correlation	Cronbach's alpha if item deleted
Q9	59.00	60.667	.555	.708
Q11	59.20	65.956	.441	.724
Q12	60.20	62.400	.631	.709
Q13	59.40	62.711	.694	.709
Q14	59.00	60.000	.599	.704
Q15	59.20	59.733	.562	.706
Q16	58.80	57.733	.571	.702
Q18	59.30	59.789	.544	.707
Q19	59.40	64.489	.236	.736
Q20	59.10	67.878	.050	.756
Q21	58.90	59.878	.575	.705
Q22	59.80	58.622	.812	.691
Q23	58.60	66.267	.363	.727
Q24	59.60	69.378	.088	.741
Q25	59.20	88.844	-.899	.818
Q26	59.70	72.456	-.219	.751
Q27	58.30	62.678	.487	.715
Q29	60.20	75.733	-.356	.772
Q31	59.60	65.156	.656	.718
Q33	59.20	70.622	.013	.742
Q34	60.50	79.833	-.747	.782
Q35	59.00	59.778	.614	.703
Q36	58.80	58.844	.794	.693

Pilot testing of the questionnaire made it possible to identify several items that demonstrated negative or extremely low correlations with the total scale score, specifically Items Q25, Q26, Q29, and Q34. Removal of these items could have increased Cronbach's alpha (to values ranging from 0.751 to 0.818), thereby improving the internal consistency coefficient of the instrument. However, it is important to consider that the questionnaire is multi-aspect in nature; therefore, exclusion of these items would be methodologically unjustified, as they are conceptually meaningful and reflect key dimensions of the construct under investigation. In addition, the small sample size renders the Cronbach's alpha coefficient unstable.

One of the methods commonly used to justify item removal or modification is item–total correlation. The values of this psychometric indicator, which reflects the extent to which an individual item is consistent with or correlated to the overall scale, ranged from –0.899 to 0.812. The vast majority of item–total correlation values were considered good and acceptable (>0.30), confirming their consistency with the overall scale. Although certain items demonstrated low or even negative correlations, their removal would be methodologically inappropriate, as these items play a conceptual role within the measured construct. This approach is consistent with the recommendations of DeVellis (2003) [12], who emphasized that conceptually important items should not be excluded solely for the purpose of increasing Cronbach's alpha.

Overall, the results of the analysis indicated that the obtained Cronbach's alpha coefficient met established reliability criteria, thereby confirming the internal consistency of the developed instrument. Based on the above considerations, a decision was made to retain these items in the final version of the questionnaire.

Finalization of the Instrument

Following the comprehensive validation process, the final version of the questionnaire was established. The finalized instrument consists of 38 items and is intended to assess the organizational aspects of uro-andrological care.

It should be noted that the informed consent form for respondents was included as part of the materials of the doctoral dissertation entitled «Organizational and Functional Model of the Male Reproductive Health Protection System» and was approved at a meeting of the Local Ethics Committee (Protocol No. 2, dated December 5, 2024).

The informed consent form provided respondents with detailed information regarding the study objectives, the number of questionnaire items, the format of questions and response options, and the approximate time required to complete the questionnaire. In addition, potential participants were informed about the voluntary nature of participation, the anonymity of the survey, and guarantees of confidentiality of the collected data. The informed consent explicitly stated that all collected information would be used exclusively for research purposes.

Thus, it can be concluded that the multi-aspect questionnaire «Organizational Aspects of Uro-Andrological Care», designed for specialists involved in organizational issues of male reproductive health, represents a validated and reliable instrument ready for use in the development and scientific substantiation of priority directions in the male reproductive health protection system.

In addition, a certificate of registration in the State Register of Rights to Copyright-Protected Objects was obtained for this questionnaire (Registration No. 62500, dated September 29, 2025).

Discussion

The development and validation of an instrument for sociological research aimed at identifying key organizational-processual and sociocultural barriers affecting the functioning of the existing uro-andrological care system represent an important step toward establishing scientifically grounded approaches to improving specialized medical care.

One of the main limitations of the present study is the absence of standardized and validated instruments that comprehensively cover all organizational aspects of uro-andrological care for the male population, which complicates direct comparison of findings with other measurement tools.

Validity and reliability constitute two fundamental components of measurement instrument evaluation. Validity refers to the extent to which an instrument measures what it is intended to measure, whereas reliability refers to the instrument's ability to produce stable and consistent results. It should be noted that reliability is closely related to validity: an instrument cannot be valid if it is not reliable; however, reliability does not depend on validity [23]. To ensure methodological rigor, a multi-stage validation procedure was conducted, including conceptual (theoretical) validity, face validity, content validity, cognitive validity (pilot study), assessment of internal consistency reliability, and finalization of the questionnaire.

Methodological emphasis during the validation process was placed on content validity, that is, qualitative validation. From the perspective of content validity, many questionnaire items demonstrated maximum values (I-CVI=1.00), indicating complete expert agreement regarding item relevance and confirming that the questions adequately reflected the intended content of the instrument. I-CVI values of 0.875 observed for certain items also exceeded the recommended threshold of 0.78. The obtained S-CVI/Ave (0.987–1.00) and S-CVI/UA (0.895–1.00) values met established standards for content validity (S-CVI/Ave≥0.90; S-CVI/UA≥0.80). Assessment of content validity represents a critical stage in the psychometric evaluation and pilot testing of sociological instruments, as it ensures alignment between questionnaire items and the stated research objectives. In other words, the content validation process involves not only quantitative but also qualitative expert assessment of each item, which is particularly important given the small sample size (n=10).

Considering the professional competence of respondents and the unambiguous wording of questionnaire items, as confirmed by the results of content validity analysis, conducting additional cognitive interviews and/or think-aloud procedures was deemed unnecessary. Moreover, the high content validity indices indicate the absence of problematic item formulations, rendering further cognitive interviewing redundant.

Given that the target population size was ≤ 50, conducting factor analysis – both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) – was considered inappropriate, as small sample sizes negatively

affect the quality and reliability of such analyses [10]. A commonly accepted rule of thumb suggests that at least 10 participants are required per item, resulting in an optimal respondent-to-item ratio of 10:1 [2], [15].

The Cronbach's alpha coefficient, reflecting the level of internal consistency among items, yielded a value of 0.739, which represents a satisfactory level and corresponds to an acceptable degree of internal reliability. It is important to emphasize that Cronbach's alpha is a property of test results obtained from a specific sample. Given that the pilot testing phase involved a small sample (n=10), test–retest reliability was not assessed. Small sample sizes directly affect the stability of Cronbach's alpha estimates. Such limitations are considered acceptable in the context of developing and validating sociological instruments. Therefore, researchers advise against relying solely on published alpha coefficients and recommend calculating alpha for each new application of the instrument [28].

During pilot testing, several items were identified as negatively correlated with the total scale score. However, it must be considered that the developed instrument is multi-aspect in nature, whereas the calculation of Cronbach's alpha assumes unidimensionality. Excluding conceptually important items solely for statistical reasons is regarded as a methodological error [12]. These items are conceptually essential for measuring organizational-processual and sociocultural domains and ensure adequate content representation of the construct.

Taken together, the comprehensive set of validation procedures confirms that the developed questionnaire is methodologically sound, content-relevant, and psychometrically acceptable for assessing the current state of uro-andrological care, identifying barriers, and designing organizational interventions aimed at improving healthcare services for men with reproductive health problems.

Conclusions

The use of valid and reliable instruments in sociological research is a key prerequisite for ensuring the accuracy of findings and the overall quality of research outcomes. The developed multi-aspect questionnaire «Organizational Aspects of Uro-Andrological Care», intended for specialists in healthcare organization, underwent a comprehensive, multi-level validation process, including assessment of conceptual, face, content, and cognitive validity, as well as evaluation of internal consistency reliability.

The obtained results confirm the methodological robustness and scientific applicability of the instrument for use in subsequent research, including multicenter studies in the field of male reproductive health. The questionnaire is suitable for identifying organizational barriers within the uro-andrological care system and for building an evidence base to optimize healthcare services for the male population at both regional and national levels of the Republic of Kazakhstan.

Author Contributions: *The authors equally contributed to the research and editorial work that led to the publication of this article.*

Conflict of Interest: *There was no conflict of interest to declare.*

Funding: *No funding.*

Publication Information: *The authors declare that no part of this article has been published previously or is under review elsewhere.*

Literature:

1. Baker P. [и др.]. The men's health gap: men must be included in the global health equity agenda. Bulletin of the World Health Organization. 2014. № 8 (92). С. 618–620.

2. Boateng G.O. [и др.]. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Frontiers in Public Health. 2018. (6). С. 149.

3. Boivin J., Takefman J., Braverman A. The Fertility Quality of Life (FertiQoL) tool: development and general psychometric properties. Fertility and Sterility. 2011. № 2 (96). С. 409-415.e3.

4. Bolarinwa O.A. Principles and Methods of Validity and Reliability Testing of Questionnaires Used in Social and Health Science Researches. Nigerian Postgraduate Medical Journal. 2015. № 4 (22). С. 195.

5. Calvert J.K. [и др.]. The male infertility evaluation still matters in the era of high efficacy assisted reproductive technology. Fertility and Sterility. 2022. № 1 (118). С. 34–46.

6. Chalyi M.E.C., Akhvlediani N.D.A., Kharchilava R.R.K. Male infertility. Urologiia. 2017. (2-supplement_2017). С. 4–19.

7. Connell J. [и др.]. 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: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation. 2018. № 7 (27). С. 1893–1902.

8. De Jonge C., Barratt C.L.R. The present crisis in male reproductive health: an urgent need for a political, social, and research roadmap. Andrology. 2019. № 6 (7). С. 762–768.

9. De Jonge C.J. [и др.]. Current global status of male reproductive health. Human Reproduction Open. 2024. № 2 (2024). С. hoae017.

10. De Winter J.C.F., Dodou* D., Wieringa P.A.* Exploratory Factor Analysis With Small Sample Sizes. Multivariate Behavioral Research. 2009. № 2 (44). С. 147–181.

11. Department of Medical Education, School of Medical Sciences, Universiti Sains Malaysia, MALAYSIA, Yusoff M. S. B. ABC of Content Validation and Content Validity Index Calculation. Education in Medicine Journal. 2019. № 2 (11). С. 49–54.

12. DeVellis R.F. Scale development: theory and applications: Sage Publications, 2003. 190 с.

13. Díaz-Mohedo E. [и др.]. The Spanish Version of the International Index of Erectile Function: Adaptation and Validation. International Journal of Environmental Research and Public Health. 2023. № 3 (20). С. 1830.

14. Eisenberg M.L. [и др.]. Male infertility. Nature Reviews. Disease Primers. 2023. № 1 (9). С. 49.

15. J N. Psychometric theory (3rd Ed.). 1994. С.240.

16. Jung J.H., Seo J.T. Empirical medical therapy in idiopathic male infertility: Promise or panacea? Clinical and Experimental Reproductive Medicine. 2014. № 3 (41). С. 108–114.

17. Kimberlin C.L., Winterstein A.G. Validity and reliability of measurement instruments used in research. American journal of health-system pharmacy: AJHP: official

journal of the American Society of Health-System Pharmacists. 2008. № 23 (65). С. 2276–2284.

18. Kimmins S. [и др.]. Frequency, morbidity and equity - the case for increased research on male fertility. Nature Reviews. Urology. 2024. № 2 (21). С. 102–124.

19. Liang Y. [и др.]. Global, regional, and national prevalence and trends of infertility among individuals of reproductive age (15-49 years) from 1990 to 2021, with projections to 2040. Human Reproduction (Oxford, England). 2025. № 3 (40). С. 529–544.

20. Lynn M.R. Determination and quantification of content validity. Nursing Research. 1986. № 6 (35). С. 382–385.

21. Moore C.G. [и др.]. Recommendations for Planning Pilot Studies in Clinical and Translational Research. Clinical and Translational Science. 2011. № 5 (4). С. 332–337.

22. Nassir A.M. [и др.]. Urological knowledge among primary health care physicians in Saudi Arabia. Saudi Medical Journal. 2019. № 5 (40). С. 483–489.

23. Nunnally, Jum C. Nunnally, J.C. (1978) Psychometric theory. 2nd Edition, McGraw-Hill, New York. - References - Scientific Research Publishing [Электронный ресурс]. URL: https://www.scirp.org/reference/ReferencesPapers?ReferenceID=453912 (Accessed: 06.05.2025).

24. Polit D.F., Beck C.T. The content validity index: Are you sure you know what's being reported? critique and recommendations. Research in Nursing & Health. 2006. № 5 (29). С. 489–497.

25. Rovito M.J. [и др.]. A Call for Gender-Inclusive Global Health Strategies. American Journal of Men's Health. 2017. № 6 (11). С. 1804–1808.

26. Shah R. [и др.]. Consensus and Diversity in the Management of Varicocele for Male Infertility: Results of a Global Practice Survey and Comparison with Guidelines and Recommendations. The World Journal of Men's Health. 2023. № 1 (41). С. 164–197.

27. Shi J., Mo X., Sun Z. [Content validity index in scale development]. Zhong Nan Da Xue Xue Bao. Yi Xue Ban = Journal of Central South University. Medical Sciences. 2012. № 2 (37). С. 152–155.

28. Streiner D.L. Starting at the beginning: an introduction to coefficient alpha and internal consistency. Journal of Personality Assessment. 2003. № 1 (80). С. 99–103.

29. Taber K. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Research in Science Education. 2018. (48). С. 1–24.

30. Tavakol M., Dennick R. Making sense of Cronbach's alpha. International Journal of Medical Education. 2011. (2). С. 53–55.

31. Waltz C., Bausell B. Nursing Research: Design, Statistics, and Computer Analysis 1981. С.186.

32. Wiering B., Boer D. de, Delnoij D. Patient involvement in the development of patient-reported outcome measures: The developers' perspective. BMC Health Services Research. 2017. № 1 (17). С. 635.

Information about the authors:

***Bayan Orazayeva** - 2nd year PhD student in Public Health, NCJSC «Semey Medical University», Semey, Republic of Kazakhstan; e-mail: borazayeva01@mail.ru; Cell phone: +77079876612; ORCID: https://orcid.org/0009-0005-0374-7884;

Zaituna Khismetova - Candidate of Medical Sciences, Professor, Head of the Department of Public Health, NCJSC «Semey Medical University», Semey, Republic of Kazakhstan; e-mail: zaituna.khismetova@smu.edu.kz; Cell phone: +77772582681; ORCID: https://orcid.org/0000-0001-5937-3045;

Merkhat Akkaliyev - PhD, Acting Associate Professor of the Department of Surgical Disciplines, NCJSC «Semey Medical University», Semey, Republic of Kazakhstan; e-mail: merhat.akkaliev@smu.edu.kz; Cell phone: +77771539854; ORCID: https://orcid.org/0000-0003-3122-7411.

***Corresponding author:**

Bayan Orazayeva – 2nd-year PhD student in Public Health at the NCJSC «Semey Medical University», Semey, Republic of Kazakhstan, https://orcid.org/0009-0005-0374-7884**.**

Postal code: Republic of Kazakhstan, Semey city, 071400, Abay Street 103.

E-mail: borazayeva01@mail.ru

Cell phone: +7 707 987 66 12