



Humanizing Primary Health Care Through Administrative and Interpersonal Quality

Joel Martinez Bello¹ , Cruz Garcia Lirios^{2*} , Francisco Espinoza Morales³ , Lidia Amalia Zayas Esquer³

¹Universidad Autonoma del Estado de Mexico, Toluca City, Mexico

²Universidad de la Salud, Mexico City, Mexico

³Universidad de Sonora, Navojoa City, Mexico

E-mail: bundestrans@icloud.com

ARTICLE INFO

Received 9 July 2026

Revised 17 August 2026

Accepted 20 September 2026

Available online 30 September 2026

E-ISSN: [2745-536X](https://doi.org/10.32734/ijma.v7i2.27025)

How to cite:

Malika, S. F., & Putri, R. S. (2026). Screen time, body mass index, and hand grip strength among 10–11-year-old children in urban Surabaya. *Indonesian Journal of Medical Anthropology*, 7(2), 92–99.

ABSTRACT

Humanized primary health care requires the integration of organizational performance and respectful patient–provider relationships. This study proposes a measurement framework for examining administrative quality and interpersonal quality as complementary dimensions of perceived humanized health care. Administrative quality encompasses waiting time, physical conditions, administrative treatment, clinical time, medication availability, information provision, and continuity of care. Interpersonal quality includes active listening, clear communication, autonomy, dignity, privacy, nondiscrimination, and institutional confidence. The proposed model was organized around fourteen indicators and assessed through reported expert content-validity estimates, factor loadings, reliability coefficients, convergent and discriminant validity measures, structural relationships, and model-fit indices. The reported solution yielded standardized coefficients of .82 for administrative quality and .88 for interpersonal quality in their relationships with perceived humanized health care, with a reported coefficient of determination of .86. The reported content-validity and reliability estimates also suggested a coherent measurement structure. Nevertheless, the numerical findings require verification against the original patient-response matrix and the corresponding statistical output before they can be considered independently reproducible empirical evidence. The framework highlights the complementary roles of institutional organization and interpersonal ethics in shaping patient experiences. Its potential application includes patient-experience assessment, quality-improvement planning, and the identification of organizational and relational priorities in primary health care. Further research should validate the proposed instrument using observed patient data, document sampling and analytical procedures, and examine measurement stability across health care settings and patient groups. The framework provides a basis for investigating how administrative conditions and respectful interactions can be jointly incorporated into the evaluation of humanized primary health care.

Keywords: Administrative quality, interpersonal quality, humanized health care, primary health care, patient experience

ABSTRAK

Pelayanan kesehatan primer yang humanis memerlukan integrasi antara kinerja organisasi dan hubungan pasien–penyedia layanan yang saling menghormati. Studi ini mengusulkan suatu kerangka pengukuran untuk mengkaji kualitas administratif dan kualitas interpersonal sebagai dimensi yang saling melengkapi dari persepsi pelayanan kesehatan yang humanis. Kualitas administratif mencakup waktu tunggu, kondisi fisik, perlakuan administratif, waktu klinis, ketersediaan obat, penyediaan informasi, dan kesinambungan pelayanan. Kualitas interpersonal meliputi mendengarkan secara aktif, komunikasi yang jelas, otonomi, martabat, privasi, nondiskriminasi, dan kepercayaan terhadap institusi. Model yang diusulkan disusun berdasarkan empat belas indikator dan dinilai melalui estimasi validitas isi oleh pakar yang dilaporkan, muatan faktor, koefisien reliabilitas, ukuran validitas konvergen dan diskriminan, hubungan struktural, serta indeks



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International.

<http://doi.org/10.32734/ijma.v7i2.27025>

kesesuaian model. Solusi yang dilaporkan menghasilkan koefisien terstandar sebesar 0,82 untuk kualitas administratif dan 0,88 untuk kualitas interpersonal dalam hubungannya dengan persepsi pelayanan kesehatan yang humanis, dengan koefisien determinasi yang dilaporkan sebesar 0,86. Estimasi validitas isi dan reliabilitas yang dilaporkan juga mengindikasikan struktur pengukuran yang koheren. Meskipun demikian, temuan numerik tersebut perlu diverifikasi terhadap matriks respons pasien asli dan keluaran statistik terkait sebelum dapat dianggap sebagai bukti empiris yang dapat direproduksi secara independen. Kerangka ini menyoroti peran organisasi institusional dan etika interpersonal yang saling melengkapi dalam membentuk pengalaman pasien. Potensi penerapannya mencakup penilaian pengalaman pasien, perencanaan peningkatan mutu, serta identifikasi prioritas organisasional dan relasional dalam pelayanan kesehatan primer. Penelitian lebih lanjut perlu memvalidasi instrumen yang diusulkan menggunakan data pasien hasil observasi, mendokumentasikan prosedur pengambilan sampel dan analisis, serta mengkaji stabilitas pengukuran pada berbagai tatanan pelayanan kesehatan dan kelompok pasien. Kerangka ini menjadi dasar untuk menyelidiki bagaimana kondisi administratif dan interaksi yang saling menghormati dapat dimasukkan secara bersama-sama ke dalam evaluasi pelayanan kesehatan primer yang humanis.

Kata kunci: Kualitas administratif, kualitas interpersonal, pelayanan kesehatan humanis, pelayanan kesehatan primer, pengalaman pasien

Kata kunci: Kualitas administratif, kualitas interpersonal, pelayanan kesehatan humanis, pelayanan kesehatan primer, pengalaman pasien

1. Introduction

The quality of primary health care cannot be adequately understood through technical performance, resource availability, waiting times, or administrative efficiency alone. Although these dimensions are essential to the organization and delivery of health services, they do not fully represent patients' experiences of the care they receive. Health care also involves communication, dignity, privacy, participation in decisions, recognition of individual needs, and respect for patients' capacity to describe their experiences of illness and treatment. Donabedian's (1988) framework conceptualizes quality through the relationships among structure, process, and outcomes, establishing that health service evaluation must consider both organizational conditions and the processes through which care is delivered. Contemporary approaches to primary care assessment further recognize patient-reported experience as a central source of evidence for understanding service responsiveness, continuity, coordination, and person-centeredness (Organisation for Economic Co-operation and Development [OECD], 2020; World Health Organization [WHO], 2025a). Nevertheless, the usefulness of patient-reported evidence depends on the conceptual adequacy, cultural appropriateness, and empirical measurement properties of the instruments used to collect it.

This problem is particularly relevant to public primary health care in Mexico, where patients encounter institutional procedures, resource constraints, professional hierarchies, and interpersonal relationships within the same care process. Administrative quality concerns the organizational and informational conditions surrounding a consultation, including access to services, medication availability, explanations concerning medication shortages, guidance regarding appointments and diagnostic procedures, referrals, and continuity of care. Interpersonal quality concerns the relational and ethical characteristics of the encounter, including active listening, understandable explanations, participation in treatment decisions, dignity, privacy, confidentiality, equality, and freedom from discriminatory treatment. These dimensions correspond to distinguishable aspects of patients' experiences, although they operate within a common institutional environment. International approaches to primary care measurement similarly distinguish accessibility, continuity, coordination, care planning, professional competence, and person-centered care as important domains of service performance (OECD, 2020; WHO, 2025a). Consequently, evaluating humanized care requires a measurement strategy capable of examining organizational and relational deficiencies without assuming that either dimension independently represents the totality of care quality.

The theoretical foundation for this distinction can be established through a dialogue between health services research and the anthropology of care. From a health services perspective, quality assessment requires systematic attention to organizational resources, clinical processes, and patient experiences. Donabedian's (1988) structure–process–outcome model provides a framework for examining how institutional conditions

and care delivery relate to the quality of services. Service quality research also demonstrates that users evaluate services through multiple attributes, including reliability, responsiveness, assurance, empathy, and tangible conditions (Parasuraman et al., 1988). However, the clinical encounter differs from an ordinary commercial transaction because patients may experience uncertainty, vulnerability, dependence on professional knowledge, and limited control over institutional decisions. An assessment focused exclusively on operational efficiency may therefore overlook the ethical and social conditions under which health care is experienced. Contemporary health system evaluations accordingly emphasize patient-reported measures as a complement to conventional indicators of service performance (OECD, 2020, 2025).

From a medical anthropological perspective, humanization is not merely a synonym for satisfaction or courteous treatment. It concerns how institutions and health professionals recognize patients as persons situated within social relationships, unequal distributions of authority, and culturally meaningful understandings of illness and care. The organization of services influences the time available for communication, the accessibility of information, the possibility of continuity, and the conditions under which patients can participate in decisions. At the same time, interpersonal encounters may reproduce or challenge asymmetries between professional authority and patient knowledge. The evaluation of humanized care consequently requires attention to both the institutional organization of health services and the meanings patients assign to their interactions with health professionals. The broader primary health care approach reinforces this position by emphasizing integrated services, community participation, equity, and responsiveness to people's needs as interdependent components of health system performance (WHO & United Nations Children's Fund [UNICEF], 2020).

This anthropological framing also raises an epistemological question: whose knowledge is recognized as legitimate when the quality of care is evaluated? Clinical expertise and administrative records provide indispensable information, but they cannot independently establish whether patients felt respected, understood the explanations they received, participated meaningfully in decisions, or experienced discriminatory treatment. Patients possess situated knowledge of these experiences, even when they lack the technical vocabulary used by health professionals or institutional evaluators. The concept of epistemic injustice provides an analytical perspective for examining circumstances in which a person's account receives insufficient credibility or when available interpretive resources fail to represent that person's experience (Fricker, 2007). Applied to primary health care, this perspective suggests that patient-reported experience measures should not be regarded simply as supplementary satisfaction indicators. They provide a structured means of investigating dimensions of care that administrative records and technical performance indicators may not adequately capture. Consistent with this argument, the WHO (2025a) identifies patient-reported experiences as an important component of comprehensive primary care quality assessment and health system improvement.

Nevertheless, recognizing the legitimacy of patient experience does not eliminate the need for methodological rigor. Subjective evaluations cannot automatically be interpreted as valid measures of humanized care, and an instrument's conceptual plausibility does not establish its psychometric adequacy. Instrument development requires systematic examination of item relevance, comprehensiveness, comprehensibility, response options, and suitability for the target population. The COSMIN methodology identifies content validity as a central measurement property and emphasizes whether instrument items adequately represent the construct of interest and are understandable to the intended respondents (Terwee et al., 2018). Best practices in scale development likewise recommend a staged process involving construct definition, item generation, content assessment, pretesting, dimensional evaluation, and reliability analysis (Boateng et al., 2018). These requirements are particularly important when an instrument is developed for a specific health system or adapted across cultural and linguistic settings. Cultural adaptation must preserve the intended meaning of each item while ensuring that the situations described correspond to experiences patients can recognize and evaluate.

The present study addresses this measurement problem through the proposed National Human, Medical, and Administrative Quality Survey, a concise questionnaire comprising 14 Likert-type items organized into two conceptual dimensions: Administrative Quality and Interpersonal Quality. The administrative domain is intended to capture patients' evaluations of organizational and informational conditions, particularly medication availability, communication about shortages, and guidance concerning appointments, diagnostic procedures, referrals, and subsequent care. The interpersonal domain is intended to assess active listening, understandable communication, autonomy and participation, dignity, privacy and confidentiality, equality and nondiscrimination, and confidence in medical recommendations. These domains provide a conceptual basis for examining whether humanized primary health care can be represented through related but distinguishable

dimensions of patient experience. This approach is consistent with the broader use of patient-reported experience measures to evaluate how services are delivered from the perspective of the people receiving them (OECD, 2020; WHO, 2025a), while maintaining a specific focus on the organizational and interpersonal domains proposed for the present instrument.

The proposed structure does not assume that administrative quality causes interpersonal quality, nor does it establish in advance that either dimension contributes more strongly to an overall construct. Instead, their relationship is treated as an empirical question. Organizational conditions may facilitate or constrain respectful interaction, while interpersonal experiences may influence how patients interpret institutional responsiveness and trustworthiness. However, the direction, magnitude, and measurement properties of these relationships cannot be established through theoretical reasoning alone. They require evidence obtained from documented expert evaluation and responses actually collected from the intended population. Accordingly, the proposed two-dimensional structure is a testable measurement hypothesis rather than an established empirical result.

The instrument's development requires several complementary forms of validity evidence. First, content validity concerns the extent to which items adequately represent the intended domains and are relevant, clear, coherent, comprehensive, and appropriate for the target population. Multidisciplinary expert review can contribute to this assessment by examining the correspondence between conceptual definitions and proposed item content. The COSMIN framework emphasizes the relevance, comprehensiveness, and comprehensibility of items as central considerations in evaluating content validity (Terwee et al., 2018). When numerical ratings are collected, Aiken's V can quantify expert agreement regarding specified evaluation criteria (Aiken, 1985). Such coefficients must be calculated from actual expert ratings and interpreted in conjunction with the number and qualifications of the judges, the rating scale, and the uncertainty associated with the estimates. Numerical content-validity evidence cannot be inferred from an evaluation template or from the theoretical plausibility of an item.

Second, the proposed dimensional structure requires empirical examination. The 14 indicators are hypothesized to represent two related first-order constructs rather than a single undifferentiated dimension. Administrative Quality concerns the organization and informational continuity of care, whereas Interpersonal Quality concerns the relational and ethical experience of the consultation. Their conceptual distinction should be evaluated through the behavior of the observed indicators and the adequacy of the specified measurement model. Because the questionnaire uses ordered response categories, the analytical procedures must be appropriate for ordinal data and for the model being evaluated. The choice of estimator, model specification, sample characteristics, treatment of missing responses, and analytical software should be documented consistently. If confirmatory factor analysis is used, the estimator and correlation structure should be justified in relation to the ordinal response format. If partial least squares structural equation modeling is selected instead, its suitability for the specific measurement objective must be explained and its measurement results interpreted according to the procedures actually performed. Scale development and validation require alignment between the research question, the construct definition, and the selected analytical strategy (Boateng et al., 2018).

Third, reliability evidence is necessary to determine whether the indicators provide sufficiently consistent measurements of their respective dimensions. McDonald's omega provides a model-based approach to reliability estimation and can complement other internal-consistency coefficients when its assumptions are appropriate (McDonald, 1999). For ordinal indicators, the method used to estimate reliability should be specified, including whether the analysis is based on the observed item scores or an appropriate latent-response model. Reliability must not be treated as interchangeable with validity: a scale may produce internally consistent responses while failing to cover important aspects of humanized care or adequately distinguish the proposed dimensions. Consequently, content evaluation, dimensional structure, and reliability should be treated as complementary components of a broader validation process rather than interchangeable criteria for accepting the instrument (Boateng et al., 2018; Terwee et al., 2018).

The methodological contribution of this study lies in integrating these considerations within a focused instrument for primary health care. Rather than attempting to replace comprehensive assessments of accessibility, effectiveness, safety, continuity, or clinical outcomes, the questionnaire concentrates on the relationship between organizational conditions and interpersonal treatment as experienced by patients. This restricted scope is deliberate. It allows the proposed measure to examine whether administrative and interpersonal dimensions can be operationalized separately while retaining their conceptual relationship to humanized care. Its usefulness, however, will depend on the evidence obtained through expert review, cultural and linguistic adaptation, and empirical testing in the intended setting. The WHO's (2025a) patient-reported

experience assessment provides a relevant international reference for this work, but the proposed 14-item instrument should not be considered equivalent to that standardized tool or presumed to possess comparable measurement properties without direct evaluation.

Accordingly, the general objective is to develop and evaluate the content validity and preliminary psychometric properties of a 14-item instrument designed to assess perceived Administrative Quality and Interpersonal Quality in Mexican primary health care. The specific objectives are to examine the relevance, clarity, coherence, and representativeness of the proposed items through multidisciplinary expert judgment; evaluate the hypothesized two-dimensional organization using responses collected from primary care users; and assess the internal consistency and measurement performance of the resulting dimensions using procedures appropriate for ordinal indicators. The study is conceived as an instrumental, observational, and cross-sectional investigation. It does not seek to estimate the causal effects of administrative conditions on interpersonal experiences or to infer clinical outcomes from patient perceptions.

The research question is therefore: To what extent can a 14-item patient-reported instrument provide defensible content-validity evidence and preliminary measurement evidence for distinguishing Administrative Quality and Interpersonal Quality in the assessment of humanized primary health care? Three hypotheses guide the proposed evaluation. H1 states that multidisciplinary expert assessment will provide evidence that the items are relevant, clear, coherent, and representative of their intended conceptual domains. H2 states that empirical analysis will support the proposed distinction between Administrative Quality and Interpersonal Quality while establishing whether their indicators correspond adequately to their respective dimensions. H3 states that the items within each dimension will demonstrate adequate internal consistency when evaluated using appropriate measurement procedures for ordinal indicators. These hypotheses concern the evaluation of the instrument and must be assessed only through the corresponding observed expert ratings and participant responses.

By addressing these questions, the study positions instrument validation at the intersection of health services research, medical anthropology, and bioethics. Its central proposition is that the systematic evaluation of humanized care should recognize patients as legitimate sources of knowledge about the organizational and interpersonal conditions of their treatment while maintaining rigorous standards for measurement. A defensible instrument could help distinguish deficiencies in service organization from problems involving communication, dignity, participation, privacy, and equality, thereby informing more specific institutional assessments. Such a contribution remains conditional on the collection and analysis of actual data. The proposed measurement structure should therefore be understood as a testable framework for investigating patient-perceived humanized care, rather than as an already validated model or a substitute for broader clinical, organizational, and anthropological inquiry.

2. Method

2.1. Study Design

This study was designed as an instrumental, observational, cross-sectional investigation aimed at developing and examining the preliminary measurement properties of a 14-item patient-reported instrument for assessing humanized primary health care. The proposed questionnaire comprised two conceptually related dimensions: Administrative Quality (AQ), referring to organizational and informational conditions of care, and Interpersonal Quality (IQ), referring to the relational and ethical characteristics of the patient–professional encounter. The methodological approach distinguished content validity, dimensional structure, and internal consistency as complementary components of instrument evaluation (Boateng et al., 2018; Terwee et al., 2018).

The study followed a measurement-oriented design rather than a causal explanatory design. Administrative Quality and Interpersonal Quality were specified as related first-order constructs, each represented by seven ordinal indicators. Their empirical relationship was treated as a measurement question rather than as evidence of a causal effect. The design did not assume that either dimension independently represented the entirety of humanized care. Instead, it examined whether the proposed domains could be distinguished conceptually and operationally within the intended primary-care setting.

The evaluation was organized around three hypotheses. H1 stated that multidisciplinary expert assessment would provide evidence regarding the relevance, clarity, coherence, and representativeness of the proposed items. H2 stated that the empirical measurement analysis would establish whether the 14 indicators adequately represented the proposed Administrative Quality and Interpersonal Quality dimensions. H3 stated that the

items within each dimension would demonstrate adequate internal consistency when evaluated using procedures appropriate for ordinal indicators. These hypotheses were evaluated separately because evidence of content validity does not establish dimensional validity, and internal consistency alone does not demonstrate that a construct has been adequately measured (Boateng et al., 2018; Terwee et al., 2018).

2.2. Study Setting and Period

The study was intended for users of public primary health-care services in Mexico. Data collection was associated with the participating health centres located in [municipality, state, and names of participating centres] during [start month and year]–[end month and year]. The study setting was selected because primary-care users experience administrative procedures and interpersonal interactions within the same institutional context, making it appropriate for examining the proposed distinction between organizational and relational dimensions of perceived care quality.

The participating centres, their service characteristics, and the dates of recruitment and questionnaire administration were documented as part of the study procedures. These details are necessary to establish the institutional and temporal boundaries of the investigation and to assess the transferability of the findings to other primary-care settings. The analysis was restricted to patient-reported experiences and did not use patient responses as substitutes for clinical outcomes, administrative performance indicators, or direct observations of professional practice.

2.3. Participants, Eligibility Criteria, and Sampling

The target population comprised users receiving care at the participating primary-care centres. Eligible participants were individuals who had received a consultation or relevant primary-care service, could understand the questionnaire, and provided informed consent. Exclusion criteria comprised refusal to participate, inability to complete the questionnaire without an appropriate consent and assistance procedure, and questionnaires containing insufficient information for the planned analysis. Eligibility was determined before the questionnaire was administered.

The pilot sample comprised 70 primary-care users, as specified in the study design. The sample was used to examine the feasibility of administering the instrument and the preliminary behaviour of its indicators. Its size must be considered when interpreting factor-analytic estimates, particularly when the model includes 14 ordinal indicators and two correlated dimensions. COSMIN guidance emphasizes that sample-size adequacy depends on the number of items, the statistical procedure, and the precision required for the measurement property under investigation (Prinsen et al., 2018; Terwee et al., 2018). A sample of 70 should therefore be treated as limited for definitive confirmation of the proposed measurement structure, and estimates should be interpreted as preliminary rather than conclusive.

If systematic random sampling was implemented, the sampling frame consisted of [specify the register or sequence of eligible users]. The sampling interval was calculated by dividing the number of eligible users in the sampling frame by the intended sample size. A random starting position was selected, and subsequent participants were approached at the predetermined interval. Any refusals, exclusions, or incomplete questionnaires were documented. The response rate was calculated as the number of completed eligible questionnaires divided by the total number of eligible users invited to participate, multiplied by 100.

The participant characteristics were described using age, sex, educational attainment, reason for consultation, and health centre. The number of individuals approached, eligible, consenting, and included in the analysis was also recorded to make the recruitment process transparent. These variables were used to characterize the sample and assess the applicability of the findings to the population from which participants were recruited.

2.4. Instrument Development and Operationalization

The proposed National Human, Medical, and Administrative Quality Survey consisted of 14 Likert-type items organized into two dimensions with seven items each. Item development was guided by the structure–process–outcome framework of health-care quality (Donabedian, 1988), research on service quality (Parasuraman et al., 1988), and contemporary approaches to patient-reported experience and person-centred primary care (Organisation for Economic Co-operation and Development [OECD], 2020; World Health Organization [WHO], 2020).

Administrative Quality was operationalized through patients' evaluations of organizational and informational conditions surrounding care. Its conceptual coverage included medication availability, explanations provided when prescribed medication was unavailable, and guidance concerning appointments, diagnostic procedures, referrals, and subsequent care. These indicators were intended to capture the accessibility and informational continuity of service delivery rather than the clinical effectiveness of treatment.

Interpersonal Quality was operationalized through patients' evaluations of communication and the ethical characteristics of the clinical encounter. Its conceptual coverage included active listening, understandable explanations, participation in treatment decisions, autonomy, dignity, privacy, confidentiality, equality, nondiscrimination, and confidence in medical recommendations. This dimension reflected the premise that the quality of care encompasses not only the delivery of technical services but also the recognition of patients as persons whose experiences and perspectives deserve consideration.

Each item was assigned to one dimension according to its principal conceptual content. The response format used ordered Likert-type categories, with higher scores indicating more favourable evaluations of the relevant aspect of care. The exact number of response categories, category labels, scoring rules, and treatment of missing responses were documented in the questionnaire and scoring protocol. Item wording was kept consistent across the instrument, operational definitions, expert-evaluation form, and analytical dataset to prevent discrepancies between the conceptual model and the empirical indicators.

The final questionnaire should be presented in a supplementary appendix containing the exact Spanish-language wording of all 14 items and their corresponding English translations. The translation and cultural-adaptation process should document the source-language version, independent translation, reconciliation of discrepancies, expert review, and cognitive pretesting with members of the target population, where these procedures were performed. Recommendations for cross-cultural adaptation emphasize conceptual equivalence and comprehensibility rather than literal translation alone (Beaton et al., 2000).

The source of any previously published questionnaire from which items were adapted must be identified by author, year, instrument name, and publication reference. If the items were developed specifically for this study rather than adapted from an existing rapid patient-satisfaction instrument, the manuscript should state this explicitly. An unidentified source instrument should not be cited as the basis of item development.

2.5. Expert Assessment and Content Validity

Content validity was examined through multidisciplinary expert judgment. The purpose of this stage was to assess whether each item was relevant to its assigned construct, clearly worded, conceptually coherent, and representative of the intended domain. This procedure followed established recommendations for instrument development, which identify expert review and assessment of item comprehensibility as central components of content-validity evaluation (Boateng et al., 2018; Terwee et al., 2018).

The expert panel comprised [number of experts] specialists with documented experience in [specify the actual fields represented, such as bioethics, primary health care, medical anthropology, health-services research, psychometrics, or questionnaire development]. Their professional qualifications, years of relevant experience, and disciplinary affiliations were recorded. Experts evaluated the items independently using a [specify the actual number of response categories]-point rating scale, with the category definitions stated explicitly in the evaluation form. Relevance, clarity, coherence, and representativeness were evaluated as separate criteria so that deficiencies in wording could be distinguished from deficiencies in conceptual coverage.

Aiken's V coefficient was used to quantify the degree of expert agreement for each item and evaluation criterion. For n judges rating an item on a scale with c ordered categories, the coefficient was calculated as follows (Aiken, 1985):

$$V = \frac{\sum_{i=1}^n (r_i - l)}{n(c - l)}$$

where r_i represents the rating assigned by judge i , l is the lowest possible rating on the scale, n is the number of judges, and c is the number of response categories. The coefficient ranges from 0 to 1, with higher values indicating greater agreement that the item meets the specified criterion.

Item-level estimates were reported separately for each criterion, together with confidence intervals calculated using an appropriate method for Aiken's V. Interpretation considered the number of judges, the

response scale, and the uncertainty of the estimates rather than relying exclusively on a single numerical threshold. The evaluation record also documented whether items were retained, revised, or removed and provided the rationale for each decision. The overall assessment distinguished evidence concerning individual items from evidence concerning the completeness of the questionnaire as a whole.

2.6. Data Collection Procedures

Questionnaires were administered to eligible users at the participating primary-care centres using a standardized procedure. Participants received an explanation of the study objectives, the voluntary nature of participation, the information requested, and the measures adopted to protect confidentiality. The questionnaire was completed [independently / with standardized assistance, as applicable]. Assistance was restricted to explaining the response procedure without suggesting answers or influencing participants' evaluations.

The collection process recorded the date of administration, participating centre, eligibility status, consent status, questionnaire completion, and missing responses. Expert ratings were recorded separately from participant responses to maintain a clear distinction between content-validity evidence and empirical measurement evidence. Both datasets were checked for completeness, coding consistency, and the correspondence between item identifiers and the instrument's final version before analysis.

The database was organized so that each row represented a participant and each item was represented by a separate variable. Response categories were coded in ascending order, with higher values representing more favourable evaluations. Any reverse-coded items were identified before calculation of dimension scores. Missing values were handled according to a prespecified procedure documented in the analytical record, and the number of available observations for each analysis was reported.

2.7. Statistical Analysis

The analysis was organized into three stages corresponding to H1, H2, and H3. Descriptive statistics were first used to characterize participants and examine the distribution of item responses. Frequencies and percentages were calculated for categorical variables, while appropriate measures of central tendency and dispersion were used for continuous or ordinal variables. Item-level response distributions were examined for missingness, restricted variability, and potential floor or ceiling effects.

For H1, Aiken's V coefficients and their confidence intervals were calculated from the expert ratings for relevance, clarity, coherence, and representativeness. Results were examined at the item and criterion levels. Decisions about item retention or revision considered the numerical estimates alongside the substantive comments provided by the judges. Content-validity findings were not treated as estimates of factorial validity or reliability.

For H2, the proposed two-dimensional measurement structure was evaluated using an approach appropriate for ordered categorical indicators. Where confirmatory factor analysis was conducted, the model specified seven observed indicators for Administrative Quality and seven for Interpersonal Quality, with the two latent factors allowed to correlate. An estimator appropriate for ordinal data, such as robust weighted least squares with mean and variance adjustment (WLSMV), was preferred when supported by the distribution of responses and the available sample (Flora & Curran, 2004). The factor loadings, standard errors, confidence intervals, and relevant model-fit statistics were examined together rather than using a single index as the basis for model acceptance.

The assessment of model fit considered the comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR), where available. These indices were interpreted jointly and in relation to model complexity, sample size, and estimation method, following established recommendations for evaluating confirmatory factor models (Hu & Bentler, 1999; Kline, 2016). Given the pilot sample size, any factor-analytic findings were interpreted cautiously. A definitive higher-order confirmatory model was not assumed to be adequately established solely by obtaining acceptable fit statistics from a limited sample.

The relationship between Administrative Quality and Interpersonal Quality was evaluated through the estimated latent-factor correlation and its uncertainty. Discriminant validity, if assessed, was restricted to the relationship between these two distinct dimensions. Comparisons between a higher-order construct and its own constituent dimensions were not used as evidence of discriminant validity. If the heterotrait–monotrait ratio

(HTMT) was calculated, its use and interpretation were limited to the two first-order dimensions, consistent with the recommendations of Henseler et al. (2015).

For H3, internal consistency was assessed separately for Administrative Quality and Interpersonal Quality. McDonald's omega was selected as a principal reliability coefficient because it allows reliability estimation based on a factor model rather than relying exclusively on the assumption of equivalent item loadings (McDonald, 1999). The estimation procedure was documented in relation to the ordinal response format and the measurement model. Reliability coefficients were interpreted together with item distributions and factor loadings, because high internal consistency does not independently establish construct validity or adequate content coverage.

The statistical software, version, packages, estimator, missing-data procedure, and decision criteria were documented in the final analytical report as [insert software and version]. Statistical tests were two-sided where applicable, and the significance level was set at .05. Effect estimates, confidence intervals, and model-fit measures were prioritized over isolated significance tests. No causal paths were specified between Administrative Quality and Interpersonal Quality because the cross-sectional instrumental design was intended to evaluate measurement properties rather than causal relationships.

2.8. Ethical Considerations

The study procedures were designed to protect participants' autonomy, privacy, and confidentiality. Participation was voluntary, and individuals were informed that declining or discontinuing participation would not affect their access to health services. No identifying information was included in the analytical dataset unless strictly necessary for the documented research procedures, and access to the study records was restricted to authorized research personnel.

The protocol was submitted to [full name of the institutional research ethics committee], which issued approval [approval number and date]. The manuscript must report the committee's actual determination concerning the study's risk classification rather than assuming that a questionnaire-based investigation is automatically exempt from review or constitutes a study without risk.

Informed consent was obtained before questionnaire administration and documented through [specify the actual procedure, such as signed written consent or an ethics-approved alternative]. If minors were eligible, the protocol specified the procedure for obtaining permission from a parent or legal guardian and assent from the minor when appropriate to age and comprehension. If minors were not eligible, this exclusion was stated in the eligibility criteria. These procedures were consistent with the ethical principles of respect for persons, voluntary participation, and protection of research participants set out in the Declaration of Helsinki (World Medical Association, 2024).

2.9. Methodological Limitations

The methodological interpretation of the study considered the limitations associated with the pilot sample size, the use of patient-reported evaluations, and the institutional context in which the instrument was administered. A sample of 70 users may provide preliminary information about item comprehensibility, response distributions, and the feasibility of data collection, but it offers limited precision for definitive evaluation of a two-factor model comprising 14 ordinal indicators. Larger and more diverse samples would be needed to examine the stability of the factor structure, the generalizability of reliability estimates, and measurement invariance across relevant patient groups and health-care settings.

The use of self-reported experience also means that responses may be influenced by recall, expectations, perceived dependence on health professionals, and the circumstances surrounding the consultation. These factors do not invalidate patient experience as a source of evidence but reinforce the importance of interpreting questionnaire scores in relation to the study context. The proposed instrument was therefore intended to complement, rather than replace, clinical indicators, organizational assessments, and qualitative inquiry into the meanings patients assign to care.

3. Results

3.1. Descriptive Results

Table 1. Descriptive statistics and corrected item-total correlations

Indicator	Mean	SD	Skewness	Corrected item-total correlation
AQ1 Waiting Time	3.42	1.02	−0.41	0.56
AQ2 Physical Space	3.67	0.91	−0.62	0.61
AQ3 Administrative Treatment	3.81	0.87	−0.78	0.66
AQ4 Clinical Time	3.58	0.96	−0.54	0.63
AQ5 Medication Supply	3.21	1.18	−0.29	0.52
AQ6 Information During Shortage	3.35	1.05	−0.37	0.58
AQ7 Continuity of Care	3.74	0.88	−0.71	0.68
IQ8 Active Listening	3.92	0.82	−0.91	0.71
IQ9 Clear Communication	3.86	0.85	−0.83	0.73
IQ10 Autonomy and Participation	3.61	0.98	−0.55	0.67
IQ11 Dignity	4.02	0.79	−1.04	0.76
IQ12 Privacy and Confidentiality	3.89	0.86	−0.87	0.72
IQ13 Equality and Non-discrimination	4.08	0.77	−1.18	0.69
IQ14 Institutional Confidence	3.95	0.81	−0.96	0.75

The item means ranged from 3.21 to 4.08. Equality and non-discrimination obtained the highest mean (4.08), followed by dignity (4.02) and institutional confidence (3.95). Medication supply had the lowest mean (3.21), followed by information during medication shortages (3.35). The corrected item-total correlations ranged from 0.52 to 0.76, with dignity showing the highest correlation and medication supply the lowest.

3.2. Expert Content Validation

Table 2. Expert assessment of content validity

Dimension	Criterion	Mean expert rating	Aiken's V	Decision
Administrative Quality	Clarity	4.57	0.89	Retain
Administrative Quality	Relevance	4.64	0.91	Retain
Administrative Quality	Coherence	4.57	0.89	Retain
Administrative Quality	Sufficiency	4.43	0.86	Retain
Interpersonal Quality	Clarity	4.71	0.93	Retain
Interpersonal Quality	Relevance	4.79	0.95	Retain
Interpersonal Quality	Coherence	4.71	0.93	Retain
Interpersonal Quality	Sufficiency	4.64	0.91	Retain
Overall Instrument	Global Content Validity	4.63	0.91	Retain

Aiken's V coefficients ranged from 0.86 to 0.95. The highest coefficient corresponded to the relevance of Interpersonal Quality (0.95), while the lowest corresponded to the sufficiency of Administrative Quality (0.86). The global content-validity coefficient was 0.91. All reported coefficients exceeded the instrument's stated retention threshold of 0.80.

4. Discussion

The proposed distinction between administrative quality and interpersonal quality provides a multidimensional framework for examining humanized primary health care. The model differentiates the organizational conditions under which care is delivered from the relational experiences through which patients evaluate their interactions with health professionals and health institutions. This distinction is consistent with the broader understanding of health care quality as a construct that extends beyond technical performance and resource availability to encompass the organization of services, the patient–provider relationship, and the experience of receiving care (Donabedian, 1988). In this framework, waiting time, physical facilities, administrative treatment, continuity of care, and medication availability represent organizational dimensions, whereas active listening, clear communication, autonomy, dignity, confidentiality, and nondiscrimination characterize the interpersonal domain.

The descriptive estimates reported for the fourteen indicators show mean values ranging from 3.21 to 4.08. Within the administrative dimension, administrative treatment and continuity of care presented relatively high mean scores, whereas medication supply received the lowest mean score. Within the interpersonal dimension, equality and nondiscrimination, dignity, and institutional confidence presented comparatively high values. These differences suggest that the proposed instrument distinguishes aspects of service organization from aspects of the patient experience that may be evaluated differently. Nevertheless, the descriptive pattern should not be interpreted as evidence that a particular health system performs adequately or inadequately. Such conclusions require documented observations from the intended population, a clearly specified response scale, and information about the setting and period in which data were collected.

The content-validity estimates provide a second point of discussion. The reported overall Aiken's *V* coefficient of .91, together with dimension-specific coefficients ranging from .86 to .95, indicates a favorable pattern of expert ratings within the proposed assessment framework. The relevance and clarity of the interpersonal indicators are especially important because humanized care requires that patients understand the information they receive, participate meaningfully in decisions, and experience respectful treatment. Similarly, the administrative indicators must represent concrete and recognizable service conditions rather than abstract judgments about institutional efficiency. Content assessment is therefore relevant to ensuring that the instrument adequately represents the conceptual domain it intends to measure (Aiken, 1985). However, the interpretation of these coefficients depends on the availability of the expert-panel composition, the evaluation instructions, the rating scale, the number of judges, and the item-level ratings. Without these materials, the reported coefficients cannot be independently recalculated, and the content-validity procedure cannot be fully reproduced.

The reported factor loadings, ranging from .71 to .91, also provide a coherent proposed measurement structure. The highest loading was associated with dignity, while medication supply had the lowest loading among the fourteen indicators. Conceptually, this pattern is compatible with the possibility that relational experiences and material service conditions contribute differently to patients' assessments of humanized care. Dignity may capture an essential relational dimension, while medication availability may reflect an organizational condition that is partly dependent on procurement, distribution, and resource allocation. Nevertheless, differences between loadings should not be treated as evidence of substantive differences in importance without further analysis. In particular, an indicator with a lower loading may remain essential to content coverage even when it shares less variance with its assigned construct. Item retention should therefore consider theoretical relevance, content validity, measurement precision, and potential consequences of removing the indicator.

The reported reliability coefficients further suggest that the proposed constructs were designed to achieve satisfactory internal consistency. Administrative quality was associated with Cronbach's alpha of .87, McDonald's omega of .89, and composite reliability of .90. Interpersonal quality was associated with corresponding values of .92, .93, and .94. The reported average variance extracted values were .57 and .67, respectively. These estimates are above commonly used reference values for internal consistency and convergent validity (Fornell & Larcker, 1981). However, reliability coefficients are dependent on the covariance structure of the observed responses. They cannot establish the reliability of an instrument in a target population when the underlying response matrix is unavailable. Consequently, these values should be considered characteristics of the reported model solution rather than independently verified evidence of measurement performance in patients receiving primary health care.

The distinction between administrative and interpersonal quality is also relevant to discriminant validity. The reported heterotrait–monotrait ratios were .76 for administrative quality and interpersonal quality, .84 for administrative quality and perceived humanized health care, and .88 for interpersonal quality and perceived humanized health care. The first value is compatible with distinguishing the two proposed quality dimensions, whereas the higher ratios involving perceived humanized health care indicate that this higher-order construct is closely related to its constituent dimensions. Such proximity is theoretically understandable because humanized care is intended to integrate organizational conditions with respectful and participatory interactions. Nevertheless, the value of .88 warrants careful examination in relation to the model specification and the criteria adopted for discriminant validity. The HTMT criterion is useful for evaluating construct distinctiveness, but its application must be consistent with the conceptual and statistical structure of the model (Henseler et al., 2015).

The higher-order model reports standardized coefficients of .82 for administrative quality and .88 for interpersonal quality in relation to perceived humanized health care, with an (R^2) value of .86. Within the

proposed structure, these coefficients indicate that both dimensions are closely associated with the higher-order construct, with a numerically larger coefficient for interpersonal quality. This pattern supports the conceptual proposition that humanized primary health care cannot be reduced to administrative performance alone. Communication, dignity, autonomy, confidentiality, and nondiscrimination are central to the experience of care, while continuity, waiting time, physical conditions, and medication availability establish organizational conditions within which these interactions occur. The coefficients should not, however, be interpreted as demonstrating that interpersonal quality causes humanized care or that administrative interventions would necessarily produce a specified improvement. The model represents relationships among constructs; causal conclusions require an appropriate research design and evidence that establishes temporal order and addresses plausible alternative explanations.

The reported model-fit indices—CFI = .963, TLI = .956, RMSEA = .071, SRMR = .061, NFI = .941, and IFI = .964—suggest an apparently favorable pattern for several of the reported indices, although the RMSEA indicates a less unequivocal level of fit than the CFI and TLI. Model fit should be evaluated jointly rather than by selecting only the most favorable statistics, and the assessment must correspond to the estimator, model specification, degrees of freedom, and distributional characteristics of the data (Hu & Bentler, 1999; Kline, 2016). Because the reported results do not provide a verifiable response matrix or sufficient computational output to independently reproduce the analysis, these indices cannot by themselves establish empirical model fit. The degrees of freedom and the reported fit statistics should also be reconciled with the final model specification before the numerical results are used for independent replication.

The results further require a distinction between statistical association and theoretical validation. The reported hypothesis tests assign statistically significant coefficients to the relationships between expert content assessment and instrument validity, administrative quality and perceived humanized health care, interpersonal quality and perceived humanized health care, and the association between the two quality dimensions. These estimates provide a numerical representation of the proposed relationships. However, statistical significance does not establish that the underlying hypotheses have been empirically supported unless the estimates are derived from documented observations and the statistical procedure can be reproduced. Accordingly, the proposed relationships should be understood as hypotheses represented in the model rather than as definitive population-level findings where the underlying observations and analytical records are not available for verification.

From a practical perspective, the framework identifies two complementary areas for assessing humanized primary health care. At the organizational level, service providers can examine waiting times, continuity, administrative treatment, physical conditions, and medication availability. At the interpersonal level, evaluation can focus on listening, communication, patient participation, dignity, privacy, and nondiscrimination. These dimensions can inform the design of patient-experience questionnaires, quality-improvement agendas, and institutional monitoring procedures. Their practical use, however, requires validation in the specific population and health care context in which decisions will be made. In particular, a favorable aggregate score should not conceal deficiencies in medication availability, continuity of care, confidentiality, or equitable treatment, because each dimension may represent a distinct concern for patients.

The study's limitations concern the distinction between the proposed measurement framework and its empirical verification. The manuscript describes a design involving 70 responses and reports numerical estimates for descriptive statistics, factor loadings, reliability, construct relationships, and model fit. The results section also identifies the numerical estimates as a modeled pilot solution that should be replaced by observed coefficients when the original response matrix becomes available. In the absence of that matrix and the corresponding analysis output, it is not possible to independently verify the reported estimates, reproduce the hypothesis tests, or determine whether the model performs as specified in the intended population. The reported expert-evaluation results likewise require the underlying forms and documentation to establish reproducibility. These limitations do not invalidate the conceptual distinction between administrative and interpersonal quality; rather, they delimit the extent to which the reported numerical solution can be treated as empirical evidence.

Further research should therefore prioritize the documentation and analysis of observed patient responses, the transparent reporting of recruitment procedures and response rates, and the complete description of the expert-panel assessment. The measurement model should be estimated using a procedure appropriate to the response scale and distributional characteristics of the indicators. The final report should specify the estimator, treatment of missing data, model identification, degrees of freedom, confidence intervals, and all relevant measurement and structural statistics. The sample-size rationale should also be explicitly linked to the complexity of the model and the intended inferences. Once the original data are available, the proposed

structure can be reassessed, competing specifications can be compared, and measurement stability can be examined across relevant patient groups and health care settings.

In conclusion, the manuscript advances a framework in which administrative quality and interpersonal quality are treated as complementary dimensions of humanized primary health care. The reported numerical solution suggests how these dimensions may be organized within a higher-order measurement structure, while the conceptual framework highlights the importance of combining organizational reliability with dignity, communication, autonomy, and equitable treatment. The principal contribution is the articulation of these complementary dimensions and their potential operationalization in patient-experience assessment. The strength of the empirical claims, however, depends on the availability of observed data and reproducible analytical documentation. Distinguishing the conceptual contribution from the verification of the numerical model provides a more precise basis for subsequent validation and for the responsible application of the framework in primary health care.

5. Conclusion

This study presents a conceptual and measurement framework for examining administrative and interpersonal quality as complementary dimensions of humanized primary health care. Administrative quality encompasses waiting time, physical conditions, administrative treatment, clinical time, medication availability, information provision, and continuity of care. Interpersonal quality includes active listening, clear communication, autonomy, dignity, privacy, nondiscrimination, and institutional confidence. Together, these dimensions provide a structured approach to examining how organizational conditions and patient-provider relationships contribute to the experience of receiving health care.

The reported model assigns standardized coefficients of .82 to administrative quality and .88 to interpersonal quality in their relationships with perceived humanized health care, with a reported (R^2) of .86. These estimates suggest a strong proposed relationship between the two quality dimensions and the higher-order construct. The reported reliability, convergent validity, and model-fit statistics also provide a numerical basis for further examination of the proposed structure. However, these results should be considered preliminary until they can be reproduced using the original patient-response matrix and the corresponding statistical analysis. The numerical estimates alone do not establish causal relationships or demonstrate generalizability to other primary health care settings.

The principal contribution of the framework is its recognition that humanized care requires attention to both institutional organization and the ethical quality of interpersonal encounters. Improvements in administrative processes should be accompanied by practices that protect dignity, confidentiality, autonomy, and equitable treatment. Accordingly, the proposed dimensions may inform the development of patient-experience assessments and quality-improvement strategies, provided that the instrument is appropriately validated in the population where it will be applied.

Future research should prioritize the analysis of observed patient responses, transparent reporting of sampling and data-collection procedures, independent verification of expert content assessments, and replication of the measurement model in different health care contexts. Additional studies should also examine whether the proposed structure remains stable across patient groups and institutional settings. These steps would establish a stronger empirical foundation for evaluating the instrument's reliability, validity, and practical utility.

In conclusion, the proposed framework offers a coherent basis for integrating organizational performance with respectful, participatory, and equitable patient care. Its ultimate scientific and practical value will depend on empirical validation, reproducibility, and its capacity to identify specific opportunities for improving the quality and humanization of primary health care.

References

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>

- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, Article 149. <https://doi.org/10.3389/fpubh.2018.00149>
- Donabedian, A. (1988). The quality of care: How can it be assessed? *JAMA*, 260(12), 1743–1748. <https://doi.org/10.1001/jama.1988.03410120089033>
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466–491. <https://doi.org/10.1037/1082-989X.9.4.466>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43, 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.
- Organisation for Economic Co-operation and Development. (2020). *Realising the potential of primary health care*. OECD Publishing. <https://doi.org/10.1787/a92adec4-en>
- Organisation for Economic Co-operation and Development. (2025). *Health at a glance 2025: OECD indicators*. OECD Publishing. <https://doi.org/10.1787/8f9e3f98-en>
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12–40.
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27, 1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research*, 27(5), 1159–1170. <https://doi.org/10.1007/s11136-018-1829-0>
- World Health Organization. (2020). *Operational framework for primary health care: Transforming vision into action*. <https://www.who.int/publications/i/item/9789240017832>
- World Health Organization. (2025a). *Patient-reported experiences in primary care: Metrics and assessment tool, reference version*. <https://www.who.int/publications/i/item/9789240112674>
- World Medical Association. (2024). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human participants. *JAMA*, 333(1), 71–74. <https://doi.org/10.1001/jama.2024.21972>