Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Defining Inadequate Control and Clinical Candidacy for PTH Replacement in Chronic Hypoparathyroidism: A Focused Systematic Review

Carlos Esteban Builes-Montaño

esteban.builes@udea.edu.co

Hospital Pablo Tobon Uribe / Universidad de Antioquia https://orcid.org/0000-0002-2418-6159

Natalia A. Rojas-Henao

Universidad de Antioquia https://orcid.org/0000-0002-5051-9455

Systematic Review

Keywords: Chronic hypoparathyroidism, inadequate control, conventional therapy, parathyroid hormone replacement therapy, PTH replacement, clinical candidacy, patient-reported outcomes, systematic review

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11662126/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare potential competing interests as follows: The authors declare the following conflicts of interest. Natalia A. Rojas-Henao was contracted by Festina Lente for this work. Carlos E. Builes-Montaño has received consulting and speaking fees from Sanofi, Novo Nordisk, Novartis, Recordati Rare Diseases, Janssen, Chiesi, Abbott, and Boehringer Ingelheim, and is a shareholder of Festina Lente.

Abstract

Background

Chronic hypoparathyroidism is conventionally treated with calcium and activated vitamin D, but a subset of patients remains insufficiently controlled despite therapy. Definitions of inadequate control and criteria for considering parathyroid hormone (PTH) replacement vary across guidelines, consensus documents, and clinical studies. This review aimed to characterize these criteria and summarize the evidence relevant to PTH replacement in adults with chronic hypoparathyroidism.

Methods

We conducted a focused systematic review with structured narrative synthesis of normative, secondary, and primary clinical evidence. PubMed/MEDLINE and Embase were searched from inception through July 2026, supplemented by backward citation searching. Technology-assisted screening and extraction were performed within a human-in-the-loop workflow. Definitions, operational criteria, quantitative thresholds, clinical candidacy for PTH replacement, trial eligibility criteria, and treatment outcomes were extracted and synthesized descriptively. No meta-analysis, formal risk-of-bias assessment, or GRADE certainty rating was performed.

Results

Of 1,618 records identified, 164 reports were included. Inadequate control with conventional therapy was addressed in 108 reports, but only 24 provided an explicit textual definition. The most frequently represented domains were biochemical control (96/108, 88.9%), clinical manifestations (67/108, 62.0%), renal risk or complications (55/108, 50.9%), treatment burden (44/108, 40.7%), and patient-reported outcomes or quality of life (34/108, 31.5%). Quantitative thresholds were reported in 45 reports and persistence or repeated-occurrence requirements in 29, with substantial variability across sources. Fourteen reports directly addressed clinical candidacy for PTH replacement and were organized into 12 framework or guideline-version groupings. Recurrent candidacy domains included symptomatic or unstable calcium control, phosphate abnormalities, renal complications, high treatment requirements, impaired quality of life, malabsorption, and treatment intolerance. Sixty-seven primary reports evaluated PTH-based treatment outcomes. Across study programs, PTH replacement reduced conventional calcium and active vitamin D requirements and showed favorable biochemical, renal, and patient-reported outcomes in selected populations, although long-term comparative evidence and head-to-head data across formulations remained limited.

Conclusions

Inadequate control of chronic hypoparathyroidism is consistently represented as a multidimensional construct, but no single definition, threshold set, persistence requirement, or eligibility algorithm is universally accepted. This review provides an evidence framework to inform subsequent structured expert deliberation for a Colombian position statement, while preserving areas of uncertainty for explicit panel judgment.

Introduction

Chronic hypoparathyroidism is a rare endocrine disorder characterized by insufficient parathyroid hormone (PTH) secretion, resulting in impaired calcium and phosphate homeostasis and a lifelong requirement for clinical and biochemical monitoring. Although most cases are postsurgical, the disease also encompasses autoimmune, genetic, idiopathic, and other acquired etiologies. Available epidemiological estimates vary across healthcare systems and case definitions, with reported prevalence generally in the range of approximately 6–37 cases per 100,000 population. Beyond hypocalcemia itself, chronic hypoparathyroidism is associated with renal, skeletal, neuropsychiatric, and other complications that contribute to long-term morbidity and healthcare use (1).

Conventional management is based primarily on oral calcium and activated vitamin D, supplemented when appropriate by native vitamin D, magnesium, dietary measures, or interventions to reduce hypercalciuria. This approach can correct serum calcium but does not restore the physiological actions of PTH on the kidney, bone, and phosphate metabolism. Long-term conventional therapy has been associated with hypercalciuria, nephrolithiasis, nephrocalcinosis, and impaired renal function, while symptoms and impaired quality of life may persist even in patients receiving stable treatment (2, 3). An updated systematic review of quality-of-life studies found impairment in at least one domain in most evaluated populations receiving standard treatment, reinforcing that biochemical control alone may not fully capture disease burden (2).

The emergence of PTH replacement therapy has shifted the therapeutic objective from replacement of downstream calcium and vitamin D requirements toward restoration of the deficient hormone. This transition has made the identification of patients who remain insufficiently controlled with conventional therapy increasingly important. However, the clinical meaning of “inadequate control” extends beyond a single biochemical value and may include persistent symptoms, renal complications, treatment burden, impaired quality of life, intolerance, or difficulty adhering to or absorbing conventional therapy. Patient-centered research has similarly shown that the lived experience of hypoparathyroidism includes physical, cognitive, functional, psychological, and social effects that are not readily represented by serum calcium alone (4).

This creates a practical decision problem. Numerical thresholds used as therapeutic targets, markers of inadequate control, trial-enrolment criteria, or indications for PTH replacement are not necessarily equivalent, and clinical studies, guidelines, consensus documents, and observational cohorts address different components of the decision. For a national position statement, these sources therefore need to

be disentangled before expert judgment is applied. The evidence review should identify which domains and operational criteria recur across the literature, where empirical evidence is available, and where decisions remain dependent on clinical judgment, patient priorities, feasibility, and local healthcare context rather than allowing published frequency alone to determine recommendations.

Accordingly, this focused systematic review aimed to identify and characterize the definitions, domains, quantitative thresholds, and persistence criteria used to describe inadequate control with conventional therapy in adults with chronic hypoparathyroidism; to distinguish clinical candidacy for PTH replacement from study-specific eligibility criteria; and to summarize the reported benefits, harms, and evidence gaps associated with PTH-based treatment. The review was conducted as the evidence-informing component of a broader initiative to develop a Colombian position statement on chronic hypoparathyroidism and to provide a structured evidence framework for subsequent expert deliberation, without itself establishing consensus criteria or a clinical decision algorithm.

Methods

Review design and objectives

We conducted a focused systematic review with structured narrative synthesis to identify and characterize the definitions, domains, thresholds, and clinical circumstances used to describe inadequate control with conventional therapy in adults with chronic hypoparathyroidism; to identify criteria used to support consideration of parathyroid hormone (PTH) replacement therapy; and to summarize reported clinical outcomes associated with PTH-based treatment.

The review was developed as the evidence-informing component of a broader initiative to develop a Colombian position statement on chronic hypoparathyroidism. The systematic review and the subsequent expert-consensus process were treated as methodologically distinct phases. Accordingly, the review was designed to identify existing definitions and recommendations, characterize empirical evidence, describe areas of consistency and heterogeneity, and identify unresolved operational questions for subsequent expert consideration. The review itself did not establish consensus criteria, generate new clinical thresholds, or determine a clinical decision algorithm.

Because substantial heterogeneity was anticipated across evidence types, study designs, definitions, interventions, outcomes, and follow-up periods, a structured narrative synthesis was selected as the principal analytical approach. Two complementary evidence streams were considered: normative and secondary evidence, and primary clinical evidence.

Reporting of the review was guided by the PRISMA 2020 statement and reporting of the literature search by PRISMA-S (5).

Eligibility criteria

Eligible sources involved adults aged 18 years or older with chronic hypoparathyroidism, irrespective of sex, geographic location, healthcare setting, or underlying etiology. Eligible etiologies included postsurgical, autoimmune, genetic or syndromic, idiopathic, infiltrative, destructive, and other acquired forms associated with insufficient PTH secretion.

Chronic hypoparathyroidism was generally considered to represent persistent hypocalcemia accompanied by an inappropriately low or normal PTH concentration. Contemporary definitions of chronic postsurgical hypoparathyroidism based on persistence beyond 12 months were recognized; however, studies using earlier definitions of permanent or chronic postsurgical hypoparathyroidism, including persistence for 6 months or longer, were retained to avoid exclusion of relevant historical evidence. The definition and duration used by each source were extracted and considered during interpretation.

Studies including mixed adult and pediatric populations were eligible when information relevant to adults could be separately identified. Pregnancy and lactation were treated as special populations and were not used to derive general eligibility criteria for PTH replacement unless findings relevant to nonpregnant adults could be distinguished.

Excluded populations included transient hypoparathyroidism, isolated acute postoperative hypocalcemia without established chronic hypoparathyroidism, pseudohypoparathyroidism and other disorders characterized primarily by PTH resistance, hypocalcemia attributable exclusively to other causes, exclusively pediatric populations, and preclinical studies. Individual case reports were excluded.

Conventional therapy and inadequate control

For this review, conventional therapy was considered management based primarily on activated vitamin D analogues, adequate dietary calcium intake, and oral calcium supplementation when required. Depending on clinical circumstances, conventional management could also include native vitamin D supplementation, magnesium replacement, dietary modification, thiazide diuretics, and other measures directed at hypercalciuria or biochemical abnormalities.

No minimum dose or duration of calcium or activated vitamin D was imposed as a review-level eligibility criterion. Instead, the review extracted the components and intensity of conventional therapy reported by each source and whether treatment was characterized as optimized, maximal, stable, standard, or otherwise adequate before inadequate control was determined.

The review did not impose a single a priori operational definition of inadequate control, unsatisfactory response, or conventional-treatment failure. Sources were eligible when they reported definitions, criteria, indicators, consequences, or patient characteristics potentially relevant to determining whether chronic hypoparathyroidism was adequately controlled.

Definitions and criteria were retained in their source-specific context and subsequently organized within a multidimensional analytical framework.

PTH replacement therapy and comparators

The intervention of interest was PTH replacement therapy or another PTH-based treatment explicitly intended to restore physiological PTH activity. Eligible interventions included recombinant human PTH(1–84), PTH(1–34)/teriparatide used as replacement therapy, palopegteriparatide/TransCon PTH, and other long-acting, sustained-release, pump-delivered, or investigational PTH formulations.

Studies were eligible irrespective of dose, route, administration schedule, treatment duration, regulatory status, or concomitant use of calcium and vitamin D.

Eligible comparators included conventional therapy with calcium and activated vitamin D, placebo plus conventional therapy, another PTH formulation or administration regimen, baseline status in single-arm or before-and-after studies, and external or historical comparator populations. Absence of an explicit comparator did not result in exclusion when a report contributed relevant evidence concerning inadequate control, treatment candidacy, treatment burden, patient characteristics, outcomes, or safety.

Non-PTH interventions were not included in the principal PTH replacement synthesis.

Evidence sources and study designs

The normative and secondary evidence stream included clinical practice guidelines, recommendations from professional societies, consensus statements, position statements, systematic reviews and meta-analyses, scoping reviews, and relevant narrative or clinical reviews.

The primary clinical evidence stream included randomized controlled trials, nonrandomized interventional studies, prospective and retrospective cohort studies, case-control studies, cross-sectional studies, before-and-after studies, and treatment-extension studies.

Narrative and clinical reviews were retained in the final corpus when they contained directly extractable information relevant to the review questions, including definitions, terminology, thresholds, treatment criteria, or contextual clinical information. They were classified separately from systematic reviews and normative documents and were not interpreted as independent empirical confirmation of a criterion or treatment effect.

No restrictions based on country or healthcare setting were applied.

Information sources and search strategy

Electronic literature searches were conducted in PubMed/MEDLINE and Embase.com from database inception through July 2026.

Three complementary search components were used:

1. chronic hypoparathyroidism and inadequate control with conventional therapy;
2. chronic hypoparathyroidism and PTH replacement therapy; and
3. chronic hypoparathyroidism and clinical guidelines, consensus documents, or recommendations.

The searches combined controlled vocabulary with free-text terminology. Search terms addressed chronic hypoparathyroidism, inadequate or suboptimal control, conventional-treatment failure or intolerance, and PTH replacement. Intervention terminology included generic, historical, developmental, and brand terminology, including PTH(1–84), rhPTH(1–84), PTH(1–34), teriparatide, palopegteriparatide, TransCon PTH, Natpara/Natpar, Yorvipath, eneboparatide, and AZP-3601.

No restrictions based on publication year or language were applied. Age and study-design filters were not used because they could exclude incompletely indexed or mixed-population studies. Ineligible populations and study designs were excluded during screening.

Electronic searching was supplemented by backward citation searching of relevant included guidelines, evidence syntheses, and primary studies. Additional publications belonging to identified trials or research programs were sought when necessary to establish links between reports or retrieve complementary outcome information.

The complete executed PubMed and Embase search strategies are provided in the supplementary material.

Record management and study selection

Retrieved records were deduplicated before screening using bibliographic information including title, authors, publication year, journal, DOI, PMID, and other identifying information, with manual reconciliation when required.

Title and abstract screening used a technology-assisted, human-in-the-loop workflow developed in Python and implemented in Google Colab. The workflow combined prespecified eligibility rules with a Google large language model to generate a provisional assessment of eligibility based on the title and abstract.

The technology-assisted output was used as decision support rather than as an autonomous eligibility decision. The provisional classification and supporting rationale were reviewed by a human reviewer, who retained final authority over whether a record was excluded, retained, or advanced for further assessment. Records with insufficient information or uncertain eligibility were retained for full-text review rather than excluded on the basis of the automated assessment alone.

Potentially eligible reports underwent full-text assessment against the predefined eligibility criteria. Full-text exclusions were documented using standardized exclusion reasons, and uncertain or discrepant cases underwent adjudication before the final analytic corpus was established.

The study-selection process was summarized using a PRISMA flow diagram.

Data extraction and human verification

Data extraction was based on a structured data dictionary developed from the review objectives and evidence domains.

Extraction used a technology-assisted workflow conceptually similar to the title and abstract screening process. A Python-based process incorporating predefined extraction rules and a Google large language model was used to identify potentially relevant passages and propose values for predefined extraction fields.

Application-generated outputs were treated as draft extractions rather than final data. Extracted information was subsequently reviewed against the source material by a human reviewer, who could accept, modify, reject, or supplement the proposed information. Unsupported or unverified application-generated values were not retained in the final analytic dataset.

Source-level extraction included bibliographic characteristics, publication type, geographic setting, study design, study population, sample size, hypoparathyroidism etiology, conventional treatment, PTH intervention and formulation, comparator, follow-up duration, and relevant outcomes.

For inadequate control, extraction captured terminology and explicit definitions; biochemical criteria; clinical manifestations; acute instability; renal risk or complications; treatment burden; patient-reported outcomes and quality of life; adherence and feasibility; malabsorption; treatment tolerability; healthcare utilization; quantitative thresholds; persistence, duration, or recurrence requirements; and requirements for prior optimization of conventional therapy.

For PTH replacement candidacy, the extraction process explicitly distinguished clinical candidacy criteria from study-specific enrolment criteria. Trial inclusion and exclusion requirements were treated as descriptors of the populations in which PTH therapies were studied and were not automatically interpreted as clinical recommendations.

For PTH replacement outcomes, extracted domains included biochemical control, calcium and activated vitamin D requirements, renal outcomes, symptoms, patient-reported outcomes, skeletal outcomes, acute instability, safety, healthcare utilization, and reported treatment effects.

Particular attention was given to preserving the original context of definitions and recommendations and to distinguishing treatment targets from inadequate-control criteria, baseline characteristics from

outcomes, randomized treatment phases from open-label extensions, and individual publications from their underlying study or research program.

Unclear or internally conflicting information was reviewed against the source report and, when relevant, associated publications before the final extraction was accepted.

Unit of analysis and management of multiple reports

The report was used as the bibliographic and descriptive unit of the review. Each eligible publication was retained as an included report to preserve complete source-level traceability across primary publications, secondary analyses, treatment extensions, reviews, guidelines, and consensus documents.

When multiple reports arose from the same underlying trial, cohort, guideline process, or research program, they were linked for interpretation. Linkage was based on trial or program names, population descriptions, intervention characteristics, recruitment information, follow-up periods, and explicit references to previous participation or associated publications.

Companion reports were retained separately in the source-level dataset but were not interpreted as independent replications of the underlying evidence.

For synthesis of definitions and clinical eligibility frameworks, the candidate criterion or construct was the conceptual unit of analysis. A single report could contribute several criteria, whereas the same criterion could be reported in multiple publications or related documents.

This represented an analytical refinement from the initial protocol, which anticipated the underlying study as the principal unit when handling multiple reports. The report-level descriptive approach was adopted to preserve transparency across heterogeneous evidence types, while linkage at the study, research-program, or guideline-family level was retained to avoid interpretative double counting.

Analytical classification of the evidence

Because terminology for publication type and study design varied across sources, reports were assigned an analytical classification for synthesis.

Primary research reports were classified as randomized interventional, nonrandomized interventional, or observational/other primary research. Secondary evidence was classified as systematic review/meta-analysis, scoping review, or narrative/clinical review. Normative evidence was classified as clinical guideline/recommendation or consensus/position statement.

These categories were used to organize and describe the evidence and were not intended to represent a hierarchy of methodological quality.

PTH interventions were additionally classified as rhPTH/PTH(1–84), PTH(1–34)/teriparatide, palopegteriparatide/TransCon PTH, or other/unspecified PTH formulations. Categories were not

mutually exclusive when a report addressed more than one formulation.

Synthesis of definitions and criteria for inadequate control

Definitions and criteria were initially retained in their source-specific wording and subsequently classified using a combined deductive and inductive approach.

The final analytical domains comprised biochemical control, clinical manifestations, renal risk or complications, treatment burden, patient-reported outcomes or quality of life, adherence or feasibility, acute instability, tolerability, healthcare utilization, and other relevant criteria. Domains were not mutually exclusive.

The synthesis examined whether each report included an explicit textual definition, one or more operational criteria, quantitative thresholds or treatment targets, persistence or recurrence requirements, requirements for prior optimization of conventional therapy, assessment of adherence, or evaluation of malabsorption.

Criteria referring to conceptually similar constructs were grouped for descriptive synthesis while preserving the terminology and operational details reported by the original sources. Terms such as inadequate control, suboptimal control, refractory disease, unsatisfactory response, and failure of conventional therapy were therefore not assumed to be fully interchangeable.

Quantitative values were retained according to their source-specific purpose. Similar numerical values used as therapeutic targets, inadequate-control criteria, eligibility criteria, or reference limits were not treated as interchangeable.

Frequency of reporting was used descriptively and was not interpreted as evidence of clinical validity, relative importance, evidentiary strength, or consensus.

Synthesis of clinical candidacy for PTH replacement

Clinical statements intended to identify patients who might be considered for PTH replacement in routine practice were analyzed separately from study-specific trial eligibility criteria.

Guidelines, recommendations, and consensus sources directly addressing clinical candidacy were grouped into framework or guideline-version families when several publications arose from the same underlying recommendation process.

Clinical candidacy criteria were synthesized according to recurrent domains including symptomatic or unstable calcium control, abnormalities of phosphate homeostasis, renal risk or complications, treatment burden, patient-reported outcomes or quality of life, malabsorption, treatment intolerance, adherence or feasibility, and prior optimization of conventional therapy.

Quantitative thresholds, persistence requirements, and definitions of optimized conventional therapy were retained in their original source context rather than converted into a single review-derived eligibility

algorithm.

Related publications arising from the same guideline or consensus framework were retained for traceability but were not treated as independent endorsements of the same recommendation.

Synthesis of PTH replacement outcomes

PTH replacement outcomes were synthesized using structured narrative methods because of substantial heterogeneity in PTH formulation, study design, population, comparator, outcome definition, and duration of follow-up.

Randomized controlled phases, nonrandomized intervention studies, open-label extensions, uncontrolled longitudinal studies, and observational analyses using historical or external comparator populations were treated as methodologically distinct forms of evidence.

When several reports arose from the same trial or research program, publications were linked and interpreted jointly to avoid treating companion reports as independent studies.

Reported estimates were retained in their original form, including proportions, changes from baseline, between-group differences, hazard ratios, and confidence intervals where available.

No meta-analysis was conducted. Statistical pooling was considered inappropriate because sufficiently homogeneous groups of studies were not identified across intervention, comparator, outcome definition, follow-up, and design, and because the principal objective of the review was to characterize operational criteria, clinical candidacy frameworks, treatment outcomes, and unresolved questions relevant to subsequent expert deliberation.

Descriptive analysis

Categorical characteristics were summarized using counts and percentages. Continuous characteristics were summarized using medians and ranges when appropriate.

Because individual reports could contribute to more than one analytical domain, domain frequencies were not mutually exclusive.

Counts represented the number of included reports in which a domain or criterion was identified and were not interpreted as measures of clinical importance or evidentiary strength.

Similarly, publication counts were not interpreted as numbers of independent studies when several reports arose from the same trial, cohort, guideline framework, or research program.

Methodological appraisal and certainty of evidence

Formal study-level risk-of-bias ratings, methodological-quality scores, and GRADE certainty assessments were not performed.

Although formal critical appraisal had been contemplated in the initial protocol, the review was subsequently focused on identifying and characterizing definitions, operational criteria, treatment-eligibility frameworks, treatment outcomes, heterogeneity, and unresolved questions to inform a subsequent structured expert-consensus process. Given the highly heterogeneous corpus and the primarily evidence-mapping and decision-informing purpose of the review, formal quality scores were not used to exclude, rank, or weight sources.

Instead, source type and study design were preserved throughout the synthesis, and clinically relevant design limitations—such as lack of concurrent comparators, historical controls, overlapping populations, limited sample size, and duration of follow-up—were considered descriptively when interpreting findings. Limitations explicitly reported by study authors were also retained during extraction.

No certainty-of-evidence ratings were assigned, and the frequency with which a criterion or recommendation appeared in the literature was not interpreted as a surrogate for methodological quality or certainty.

Consequently, review conclusions were framed as a descriptive synthesis of the available evidence and as identification of unresolved decision points rather than as graded clinical recommendations.

Source-level traceability

Individual publications were cited in the main-text narrative when a source-specific definition, recommendation, quantitative threshold, or outcome estimate was described.

Reports contributing only to aggregate descriptive or thematic synthesis were not necessarily cited individually in the main text. All included reports were retained in source-level supplementary tables to preserve traceability across the complete review corpus.

Thus, the number of included reports was not expected to equal the number of publications explicitly cited in the narrative.

Protocol, amendments, and deviations

A review protocol was developed before completion of the evidence synthesis but was not prospectively registered. The review was undertaken primarily as the evidence-informing component of a broader Colombian position-statement initiative.

Several methodological refinements were introduced during review conduct. The report, rather than the underlying study, was used as the bibliographic and descriptive unit to preserve source-level traceability, while companion publications were linked into study, research-program, or guideline families for interpretation. Narrative and clinical reviews containing directly extractable information were retained as a distinct secondary-evidence category. Technology-assisted screening and extraction workflows were implemented using Python-based scripts and large language models with human final decision authority.

No meta-analysis was undertaken because of the clinical and methodological heterogeneity of the evidence. Formal risk-of-bias and certainty-of-evidence assessments contemplated in the original protocol were not performed because the final review was focused on descriptive characterization of definitions, clinical criteria, candidacy frameworks, treatment outcomes, and evidence gaps for subsequent expert deliberation.

During final dataset reconciliation, one individual case report that had remained in an intermediate working dataset was excluded in accordance with the eligibility criteria.

These changes were made explicit in the completed review and were not used retrospectively to modify criteria for the purpose of improving or altering the direction of the findings.

Ethics

The review used data from published or otherwise publicly accessible sources and did not involve recruitment of participants, intervention in patient care, or collection of identifiable individual-level patient information. Research ethics committee approval and individual informed consent were therefore not required for the systematic review component.

The subsequent expert-consensus process was considered a separate methodological phase and is not reported as part of this systematic review.

Results

Study selection

The literature search identified 1,618 records. After removal of 550 duplicates, 1,068 records were screened by title and abstract, of which 746 were excluded. The remaining 322 reports were sought for retrieval; all were obtained and assessed in full text. Following full-text assessment and final adjudication, 158 reports were excluded and 164 reports were included in the review (Fig. 1).

Because the included evidence comprised primary publications, extension studies, secondary analyses, guidelines, consensus documents, and multiple reports arising from the same study or research program, the 164 included records were treated as reports rather than as 164 independent studies.

Characteristics of the included evidence

The 164 included reports were published between 1975 and 2026, with a median publication year of 2020. Eighty-four reports (51.2%) were published between 2020 and 2026, 53 (32.3%) between 2015 and 2019, and 27 (16.5%) before 2015.

For descriptive purposes, and because the original publication and study-design labels were heterogeneous, reports were analytically grouped into 97 primary research reports, including 21

randomized interventional reports, 31 nonrandomized interventional reports, and 45 observational or other primary research reports. The remaining evidence comprised 46 reviews or evidence syntheses, 20 normative or consensus documents, and one modeling/pharmacology report (Table 1).

Among the 46 review or evidence-synthesis reports, 34 were narrative or clinical reviews, 11 were systematic reviews or meta-analyses, and one was a scoping review. Normative sources comprised 13 clinical guidelines or recommendations and seven consensus or position statements.

Forty reports were multinational or multiregional, whereas 124 were conducted in a single location. Among single-location reports, the United States contributed the largest number (n = 47), followed by Italy (n = 14), Denmark (n = 8), and Spain (n = 5).

Sample sizes varied substantially across primary research reports. Randomized interventional reports included a median of 62 participants (range, 8–192), nonrandomized interventional reports a median of 27 participants (range, 5–123), and observational or other primary reports a median of 120 participants (range, 2–12,424). Sample sizes were not pooled across reports because several publications represented extensions, secondary analyses, or overlapping populations from the same underlying study or research program.

Among the 52 primary interventional reports, rhPTH/PTH(1–84) was evaluated or explicitly addressed in 28 reports, PTH(1–34)/teriparatide in 20, and palopegteriparatide/TransCon PTH in 11. These categories were not mutually exclusive because some reports evaluated or discussed more than one PTH formulation. Characteristics of all included reports are provided in Supplementary Table S1, and their temporal distribution according to evidence type is shown in Supplementary Figure S1.

Definitions and criteria for inadequate control with conventional therapy

Of the 164 included reports, 108 (65.9%) addressed the concept of inadequate control with conventional therapy. Terminology varied across the literature and included *inadequate control*, *not adequately controlled*, *poor control*, *refractory disease*, *suboptimal response*, *failure of conventional therapy*, and intolerance or limitations of conventional treatment. Although explicit terminology was identified in 107 of the 108 reports, only 24 (22.2%) provided an explicit textual definition of inadequate control. Eighty-six reports (79.6%) described one or more criteria used to characterize inadequate control, 45 (41.7%) included quantitative thresholds or treatment targets, and 29 (26.9%) incorporated some requirement related to persistence, duration, or repeated occurrence. These definitional and operational characteristics are summarized in Table 2, with source-level explicit definitions and quantitative criteria provided in Supplementary Tables S2 and S3.

Inadequate control was generally conceptualized as multidimensional rather than solely as failure to achieve a biochemical target (Fig. 2). Biochemical control was the most frequently represented domain (96/108, 88.9%), followed by clinical manifestations (67/108, 62.0%), renal risk or complications (55/108,

50.9%), treatment burden (44/108, 40.7%), and patient-reported outcomes or quality of life (34/108, 31.5%). Adherence or feasibility (13/108, 12.0%), acute instability (9/108, 8.3%), tolerability (7/108, 6.5%), and healthcare utilization (5/108, 4.6%) were less frequently incorporated. These domains were not mutually exclusive, and individual reports commonly combined biochemical abnormalities with persistent symptoms, renal complications, high treatment requirements, or impaired quality of life.

The operationalization of inadequate control evolved over time and differed across frameworks. The 2019 European Delphi panels characterized not adequately controlled disease using repeated low albumin-adjusted serum calcium, broad fluctuations in serum calcium, hyperphosphatemia, hypercalciuria, persistent symptoms or complications, and high requirements for calcium or active vitamin D (6). A subsequent Spanish study applied a classification combining clinical chemistry with patient wellness, distinguishing patients according to the presence or absence of biochemical abnormalities and symptoms, comorbidities, or complications (7).

In 2022, the International Task Force/Second International Workshop framework defined inadequate control as the presence of any of the following despite conventional therapy: symptomatic hypocalcemia, hyperphosphatemia, renal insufficiency, hypercalciuria, or poor quality of life (8–10). Related recommendations additionally identified poor adherence, malabsorption, intolerance to large doses of calcium or active vitamin D, and high conventional-therapy requirements as circumstances in which PTH replacement might be considered. Because these publications arose from the same guideline framework, they were interpreted as related reports rather than as independent definitions.

More recent guidance provided greater operational detail. The 2025 European Society of Endocrinology guideline identified frequent calcium fluctuations requiring at least three treatment adjustments per year, symptomatic hypocalcemia, impaired quality of life attributable to chronic hypoparathyroidism, reduced kidney function with an eGFR < 60 mL/min/1.73 m², hypercalciuria, and hyperphosphatemia as features of inadequate control despite optimized conventional treatment (11). The 2025 Hellenic recommendations incorporated a minimum period of 3–6 months of intensified conventional therapy and specified serum calcium < 8.0 mg/dL, or < 8.2 mg/dL in the presence of symptoms; urinary calcium > 300 mg/day in men or > 250 mg/day in women; eGFR < 60 mL/min or an annual eGFR decline > 5 mL/min; serum phosphate ≥ 5.5 mg/dL; impaired quality of life; malabsorption; and intolerance to high-dose calcium or alfacalcidol (12). Earlier Hellenic recommendations required either persistent hypocalcemia despite high-dose conventional therapy or the presence of at least two complication-related criteria for more than 3 months (13).

Across the evidence base, however, no single quantitative threshold or persistence criterion was applied consistently. Serum calcium thresholds most commonly clustered around 8.0 mg/dL (2.0 mmol/L), whereas thresholds for serum phosphate, calcium-phosphate product, urinary calcium, kidney function, and calcium or active vitamin D requirements varied across reports. Some sources used these values as explicit indicators of inadequate control, whereas others used similar values as therapeutic targets,

eligibility criteria, or reference limits. Consequently, these thresholds were not considered interchangeable.

Requirements preceding classification as inadequately controlled were also inconsistently specified. Optimization of conventional therapy was explicitly required in 34 of 108 reports, assessment of adherence in 10, and evaluation of malabsorption in 17. In many other reports, these prerequisites were either not specified or were considered not applicable. Overall, the evidence therefore supported a multidimensional characterization of inadequate control but did not identify a single universally accepted definition, combination of domains, quantitative cutoff, or minimum duration of persistence.

Eligibility for PTH replacement therapy

Eligibility or selection criteria related to PTH therapy were reported in 136 included reports; however, these data represented different constructs. For synthesis, study-specific enrolment criteria were distinguished from clinical criteria used to identify patients who might be considered for PTH replacement in routine care. Fourteen reports directly formulated or discussed clinical candidacy for PTH replacement and were organized into 12 distinct framework or guideline-version groupings. These clinical candidacy frameworks are summarized in Table 3, with source-level statements, thresholds, persistence requirements, and optimization criteria provided in Supplementary Table S4. The remaining eligibility information largely described populations enrolled in clinical studies or contextualized recommendations from other sources and was not interpreted as defining clinical candidacy. The evolution of the principal candidacy domains across guideline frameworks is shown in Fig. 3.

Across clinical guidance and consensus sources, PTH replacement was generally positioned as an option for adults with chronic hypoparathyroidism who remained inadequately controlled despite conventional therapy, rather than as routine first-line treatment. Earlier guidance framed candidacy primarily in terms of inability to control hypocalcemia with calcium and active vitamin D (14), whereas subsequent frameworks progressively incorporated biochemical instability, renal risk, treatment burden, and patient-reported outcomes (8, 11, 15, 16).

The 2022 International Task Force/Second International Workshop framework represented a broader multidimensional approach to candidacy. PTH replacement was proposed for patients who were not adequately controlled on conventional therapy, with symptomatic hypocalcemia, hyperphosphatemia, renal insufficiency, hypercalciuria, or poor quality of life identified as indicators of inadequate control. Poor adherence, malabsorption, intolerance to large doses of calcium or active vitamin D, and high conventional-treatment requirements were additionally described as circumstances in which PTH replacement could be considered (8–10).

Renal abnormalities were recurrent candidacy features across later frameworks. These included hypercalciuria, reduced kidney function, nephrolithiasis, nephrocalcinosis, and progressive deterioration in renal function. The 2025 European Society of Endocrinology guideline included an eGFR < 60 mL/min/1.73 m² and urinary calcium excretion > 300 mg/24 h in men or > 250 mg/24 h in women, or > 4

mg/kg/24 h, among features supporting consideration of PTH replacement (11). Similar thresholds appeared in Hellenic recommendations, which additionally incorporated nephrolithiasis or nephrocalcinosis and an annual eGFR decline > 5 mL/min (12). Earlier Hellenic guidance used a more restrictive construct in asymptomatic patients, requiring at least two complication-related criteria to persist for more than 3 months (13).

Treatment burden also emerged as a recurrent consideration. Guidance documents referred to high requirements for oral calcium and active vitamin D, excessive medication burden, or inability to tolerate the doses required to achieve biochemical control. Quantitative criteria varied, with calcium requirements generally in the range of > 2 to 3 g/day and active vitamin D or analogue requirements varying according to the preparation and framework (12, 13, 16). Malabsorption and gastrointestinal disorders affecting calcium or vitamin D absorption were also incorporated into several frameworks (8, 16, 17).

Patient symptoms and quality of life increasingly contributed to candidacy in addition to biochemical measures of control. The International Task Force/Second International Workshop explicitly included poor quality of life among the criteria for inadequate control (8, 9). The 2025 European Society of Endocrinology guideline similarly included impaired quality of life attributable to chronic hypoparathyroidism, together with symptomatic hypocalcemia and frequent calcium fluctuations (11). The 2025 Hellenic recommendations provided a more explicit operational example, incorporating an SF-36 score < 50 and/or impairment identified with a hypoparathyroidism-specific patient-reported outcome instrument (12). No common patient-reported outcome threshold was used across frameworks.

Prior optimization of conventional therapy was an explicit component of several candidacy frameworks, although its definition was not uniform. Formulations ranged from failure despite maximal effort or optimized conventional treatment (8, 11) to at least 3–6 months of intensified treatment (12). Earlier Hellenic recommendations referred to maximum doses of calcium and alfacalcidol together with persistence of biochemical or complication-related criteria for more than 3 months (13). Thus, neither the required duration of inadequate control nor what constituted an adequate therapeutic trial was standardized across sources.

Clinical trial eligibility criteria were substantially more restrictive than clinical candidacy criteria. Primary studies frequently specified age, minimum disease duration, baseline calcium concentrations, minimum supplement requirements, renal function, prior PTH exposure, concomitant medications, and extensive safety exclusions. These requirements were treated as descriptors of the populations in which PTH therapies were studied rather than as criteria that patients must satisfy to be considered for PTH replacement in clinical practice. Study-specific inclusion and exclusion criteria are therefore reported separately in Supplementary Table S5.

Overall, clinical candidacy for PTH replacement was consistently linked to failure or limitations of conventional therapy, but the specific combination, number, magnitude, and persistence of qualifying features varied across guidance documents. The recurrent domains were symptomatic or unstable

calcium control, abnormalities of phosphate homeostasis, renal risk or complications, high treatment requirements, impaired quality of life, and factors limiting the feasibility or tolerability of conventional therapy. No single eligibility algorithm was shared across the included frameworks.

Effects of PTH replacement therapy

Sixty-seven primary reports directly evaluated clinical, biochemical, renal, patient-reported, skeletal, treatment-burden, or safety outcomes during exposure to a PTH-based intervention. These reports were not interpreted as 67 independent studies because several arose from the same randomized trials, open-label extensions, or longitudinal cohorts. Major publication families included REPLACE/RELAY and its extensions, PaTH Forward, PaTHway, PARALLAX and its extension, a Danish randomized PTH(1–84) program, the Columbia longitudinal rhPTH(1–84) cohort, and longitudinal PTH(1–34) programs. Their principal outcome domains are summarized in Table 4, and the distribution of evidence across formulations, designs, and outcome domains is shown in Fig. 4. Study–report linkage is provided in Supplementary Table S6, and report-level outcome data are presented in Supplementary Table S7.

Biochemical control and reduction of conventional therapy

In REPLACE, the composite efficacy endpoint—reduction in oral calcium and active vitamin D while maintaining serum calcium—was achieved by 48 of 90 patients (53%) receiving rhPTH(1–84) compared with 1 of 44 (2%) receiving placebo (18). A secondary analysis from the same trial showed greater reductions in serum phosphate and the calcium-phosphate product with rhPTH(1–84) than with placebo (19). In long-term extensions of the REPLACE/RELAY program, serum calcium remained within the target range while oral calcium and active vitamin D requirements remained substantially reduced through 5–6 years; at month 60, mean oral calcium and calcitriol doses had fallen by 53.4% and 75.7%, respectively (20, 21).

With palopegteriparatide, 91% of participants in PaTH Forward were independent of conventional therapy by week 26 (22). In the phase 3 PaTHway trial, 79% of participants receiving palopegteriparatide met the multicomponent primary efficacy endpoint at week 26 compared with 5% receiving placebo, and 93% achieved independence from conventional therapy (23). At week 52, 95% of participants were independent of conventional therapy and none required active vitamin D (24).

Renal outcomes

Renal findings varied by treatment, duration, and study design. In the 24-week REPLACE trial, 24-hour urinary calcium decreased in both the rhPTH(1–84) and placebo groups, with no significant between-group difference (18). During longer-term rhPTH(1–84) treatment, urinary calcium declined during the first several years and eGFR remained stable in extension and longitudinal cohorts (20, 21, 25). In the longest Columbia follow-up, urinary calcium declined through year 8 but returned toward baseline during years 9–12, while eGFR remained stable (26).

With palopegteriparatide, mean 24-hour urinary calcium fell from 392 to 220 mg/day during the 26-week randomized PaTHway period, compared with a smaller change from 329 to 292 mg/day with placebo

(23). At week 52, urinary calcium remained reduced and mean eGFR had increased by 9.3 mL/min/1.73 m² from baseline; the mean eGFR change remained approximately 9.0 mL/min/1.73 m² at week 104 (27, 28). These later renal estimates were derived from open-label follow-up after the controlled phase.

Comparative observational renal and cardiovascular outcomes

Several retrospective analyses compared rhPTH(1–84)-treated participants drawn from clinical trial cohorts with historical or external controls. Across five years, treated patients had a more favorable eGFR trajectory than controls in two related analyses (29, 30). In a separate historical-control cohort, rhPTH(1–84) treatment was associated with lower risks of incident chronic kidney disease and sustained ≥ 30% eGFR decline (31). Another analysis reported fewer cardiovascular events among rhPTH(1–84)-treated patients than historical controls (32).

Because these analyses were observational, used external or historical comparators, and partly drew treated participants from overlapping clinical trial populations, the findings were interpreted as associations rather than causal treatment effects.

Symptoms, patient-reported outcomes, and quality of life

Patient-reported outcomes were heterogeneous across studies and instruments. In a 6-month randomized study of PTH(1–84), changes in SF-36v2, WHO-5, and symptom measures did not differ significantly from placebo (33). An analysis of health-related quality of life from REPLACE identified improvements from baseline in several SF-36 domains in the rhPTH(1–84) group, but overall between-group differences were not statistically significant (34).

In contrast, a later randomized trial in symptomatic adults found greater improvement with rhPTH(1–84) than placebo in the HypoPT-SD symptom subscale, FACIT-Fatigue, and the SF-36v2 physical component score (35). In the randomized phases of PaTH Forward and PaTHway, palopegteriparatide was associated with improvements in disease-specific physical, cognitive, and daily-life domains compared with placebo (22, 23). Improvements in patient-reported outcomes were subsequently described during longer-term follow-up of the PaTH Forward and PaTHway programs (24, 36).

Skeletal outcomes

PTH replacement consistently altered bone turnover, but skeletal effects were site- and time-dependent. Long-term rhPTH(1–84) treatment in the Columbia cohort was associated with increased lumbar-spine bone mineral density, relative stability or smaller changes at the hip, and reductions at the distal one-third radius (25). At follow-up extending to 12 years, cortical volumetric BMD at the tibia was lower than at earlier time points and cortical porosity increased, whereas estimated failure load remained above values observed in normal controls (26).

During three years of palopegteriparatide treatment in PaTH Forward, bone-turnover markers rose early and then partially stabilized, while BMD decreased from elevated baseline values, predominantly during the first 26 weeks; mean Z-scores remained above zero at all measured sites at week 162 (37).

Safety

Adverse events were common in both active-treatment and comparator groups of randomized trials, although most were reported as mild or moderate. In REPLACE, adverse events occurred in 93% of patients receiving rhPTH(1–84) and 100% receiving placebo, with serious adverse events in 11% and 9%, respectively (18).

In PaTH Forward, adverse events were reported in 62.7% of participants receiving palopegteriparatide, with no serious adverse event considered treatment-related and no adverse-event-related discontinuation or death during the reported period (22). In PaTHway, adverse events during the 26-week randomized phase occurred in 82% of palopegteriparatide-treated participants and 100% of placebo recipients; serious adverse events occurred in 8.2% and 14%, respectively (23). At week 52, 90% of participants exposed to palopegteriparatide had experienced a treatment-emergent adverse event and 10% a serious treatment-emergent adverse event; injection-site reactions, hypercalcemia, nausea, headache, and hypocalcemia were among treatment-related events reported in at least 5% of participants (24).

Longer-term uncontrolled studies reported fractures, nephrolithiasis, and episodes of hypo- or hypercalcemia, but their sample sizes and designs were insufficient to estimate comparative risks of uncommon or very long-term adverse outcomes.

Heterogeneity, evidence gaps, and unresolved issues

The evidence base was heterogeneous at several levels, including the definition of inadequate control, thresholds used to operationalize individual domains, requirements for prior optimization of conventional therapy, clinical candidacy for PTH replacement, outcome selection, follow-up duration, and study design (Table 5). Although 108 reports addressed inadequate control, only 24 provided an explicit textual definition, 45 reported quantitative thresholds or treatment targets, and 29 incorporated a persistence, duration, or repeated-occurrence requirement. Consequently, the same clinical domain could be represented by different numerical cutoffs, temporal requirements, or combinations of criteria across frameworks (6–8, 11, 12).

Heterogeneity was also evident in the evidence supporting PTH replacement. Sixty-seven primary reports directly evaluated PTH-based treatment or exposure, but a substantial proportion represented companion publications, secondary analyses, or extensions of the same underlying study programs. Randomized controlled programs were followed by open-label extensions in the REPLACE/RELAY and PaTHway programs (18, 20, 21, 23, 24), whereas several longer-term renal and cardiovascular analyses used historical or external comparator cohorts (29–32). Accordingly, publication counts were not

interpreted as numbers of independent studies, and estimates derived from uncontrolled extensions or historical-control analyses were not considered methodologically equivalent to randomized comparisons.

Outcome definitions and measurement approaches varied across domains. Renal outcomes ranged from urinary calcium and eGFR to nephrolithiasis, nephrocalcinosis, incident chronic kidney disease, and longitudinal eGFR decline. These surrogate and clinical outcomes were assessed over different time horizons and in different study designs, limiting direct comparison. Patient-reported outcomes were similarly heterogeneous: older randomized studies using generic instruments did not consistently identify between-group improvements, whereas later trials enrolling symptomatic patients or using disease-specific measures reported more favorable changes (22, 23, 33–35). Skeletal outcomes included bone-turnover markers, areal BMD, volumetric BMD, microarchitecture, cortical porosity, and fractures, with direction and magnitude of change dependent on skeletal site and duration of treatment (25, 26, 37).

Long-term clinical outcomes remained incompletely characterized. Extended follow-up was available for selected rhPTH(1–84) and palopegteriparatide programs, but many reports identified limited sample size, lack of concurrent controls, or insufficient duration for uncommon outcomes. Safety data were therefore more informative for frequent treatment-emergent events than for rare or very long-term complications. Likewise, comparative evidence across PTH formulations was limited, and the review did not identify a broad set of head-to-head studies capable of establishing relative effects across rhPTH(1–84), PTH(1–34)/teriparatide, and palopegteriparatide. The distribution of these evidence gaps across key decision domains is summarized in Fig. 5, with report-level gaps and limitations provided in Supplementary Table S8.

Several operational questions therefore remained unresolved by the literature alone. These included the minimum combination of domains required to define inadequate control, the duration or frequency with which abnormalities should persist, what constitutes adequately optimized conventional therapy, which biochemical or renal thresholds should trigger reconsideration of treatment strategy, how treatment burden and patient-reported outcomes should be operationalized, and how evidence should be extrapolated across PTH formulations and patient subgroups. These unresolved questions are summarized in Supplementary Table S9 and represent evidence-derived decision points for subsequent expert consideration rather than recommendations or consensus conclusions.

Discussion

This focused systematic review shows that inadequate control of chronic hypoparathyroidism is consistently conceptualized as a multidimensional problem but remains incompletely standardized. Although biochemical abnormalities were the most frequently represented domain, clinical manifestations, renal risk, treatment burden, patient-reported outcomes, feasibility, and treatment tolerability were recurrently incorporated across the literature. Importantly, only a minority of reports

provided an explicit definition, and quantitative thresholds, persistence requirements, and prerequisites for optimization of conventional therapy varied substantially. These findings suggest that the principal challenge is no longer identifying whether inadequate control extends beyond serum calcium but determining how multiple relevant domains should be combined and operationalized in clinical decision-making.

A similar evolution was observed for candidacy for PTH replacement therapy. Earlier frameworks focused predominantly on failure to maintain acceptable calcium control with conventional treatment, whereas later approaches increasingly incorporated renal complications, phosphate abnormalities, high supplement requirements, impaired quality of life, malabsorption, and treatment intolerance. At the same time, this review identified an important distinction between clinical candidacy and clinical-trial eligibility. Trial populations were commonly defined by restrictive biochemical, age, renal-function, supplement-dose, and safety criteria that reflect study design and regulatory requirements rather than necessarily defining which patients should receive PTH replacement in routine care. Preserving this distinction is essential to avoid converting research-enrolment criteria into clinical indications without explicit justification.

The treatment evidence provides a complementary but uneven basis for decision-making. Randomized trials most consistently demonstrated the capacity of PTH replacement to reduce or eliminate conventional calcium and active vitamin D requirements while maintaining biochemical control. More recent trials also reported favorable effects on urinary calcium and disease-specific patient-reported outcomes. However, much of the longer-term evidence regarding renal function, skeletal outcomes, cardiovascular events, and uncommon adverse outcomes derives from open-label extensions, uncontrolled longitudinal cohorts, or historical-control analyses. Furthermore, multiple publications arose from the same underlying study programs, and direct head-to-head evidence across PTH formulations remains limited. The available evidence therefore supports the clinical relevance of PTH replacement as a therapeutic strategy but does not provide a sufficient basis for assuming equivalence or comparative superiority among formulations or for defining a single evidence-derived eligibility algorithm.

The main contribution of this review to the subsequent Colombian expert process is therefore to provide a structured evidence framework rather than a predetermined consensus solution. The literature identifies recurrent domains and plausible operational thresholds, but leaves several value-sensitive or context-dependent questions unresolved: how many domains should be required, whether individual domains can independently establish inadequate control, how long abnormalities should persist, what constitutes adequately optimized conventional therapy, how patient-reported burden should be weighted, and whether criteria can be extrapolated across PTH formulations. Such questions appropriately require explicit expert judgment informed by evidence rather than retrospective selection of thresholds from literature. Decision frameworks developed for guideline panels emphasize that evidence should be considered together with clinical relevance, patient values, feasibility, resource implications, equity, and local implementation context, while maintaining transparency about the basis for each judgment (38,

39). In the subsequent consensus phase, the review should therefore function as the evidence dossier against which panel judgments are made, with the consensus process, panel composition, voting rules, disagreement, and adaptations prospectively specified and transparently reported in accordance with contemporary consensus-reporting guidance (40).

Several limitations should be considered. The protocol was developed prospectively but was not registered before review conduct, limiting independent verification of the timing of methodological decisions. Formal study-level risk-of-bias assessment and GRADE certainty ratings were not performed; consequently, treatment-effect findings should be interpreted descriptively and should not be read as graded comparative-effectiveness conclusions. The corpus intentionally combined primary studies, guidelines, consensus documents, systematic reviews, and selected narrative reviews to characterize definitions and decision frameworks, creating substantial methodological heterogeneity and preventing meaningful meta-analysis. Several study programs generated multiple companion publications, although these were explicitly linked to reduce interpretative double counting. Technology-assisted screening and extraction may introduce additional sources of error despite human verification of final decisions and extracted data. Finally, the evidence was generated largely outside Colombia, and differences in access to therapies, monitoring capacity, healthcare organization, resource constraints, and patient preferences may affect transferability to the Colombian setting.

These limitations are balanced by several strengths, including the broad capture of normative and empirical evidence, explicit separation of inadequate-control criteria from treatment candidacy and trial eligibility, preservation of source-specific thresholds rather than creation of post hoc cutoffs, linkage of companion publications, and report-level traceability across the complete review corpus. Taken together, the findings provide a reproducible map of what is supported, what varies across sources, and what remains unresolved. The next step should not be to force the heterogeneous literature into a single retrospective definition, but to use these findings to structure expert deliberation around clearly formulated clinical scenarios and prespecified decision rules. Consensus methodology has long emphasized that structured expert judgment is most useful when empirical evidence is incomplete or does not itself resolve the clinical decision, while the evidence presented to panelists and the rules governing agreement remain explicit and reproducible (40, 41).

Conclusion

Inadequate control of chronic hypoparathyroidism is consistently represented in literature as a multidimensional construct rather than a purely biochemical state. Recurrent domains include symptomatic or unstable calcium control, phosphate abnormalities, renal risk or complications, treatment burden, impaired quality of life, and limitations in the feasibility or tolerability of conventional therapy. However, literature does not support a single universally accepted definition, threshold set, persistence requirement, or clinical eligibility algorithm for PTH replacement.

PTH replacement therapy can reduce dependence on conventional calcium and active vitamin D therapy and has shown favorable biochemical, renal, and patient-reported outcomes in selected study populations, but the evidence remains heterogeneous across formulations, designs, follow-up periods, and outcome definitions. These findings should therefore be interpreted as an evidence map rather than as graded comparative-effectiveness conclusions.

The principal value of this review is to distinguish areas of recurring clinical agreement from questions that remain unresolved by the evidence alone. These evidence-derived domains, thresholds, and uncertainties provide a structured foundation for the subsequent Colombian expert panel, which should determine how they are prioritized, combined, and operationalized within the local clinical and healthcare context using prospectively defined consensus procedures.

Declarations

Commissioning body: This review was commissioned and conducted under the auspices of the Asociación Colombiana de Endocrinología, Diabetes y Metabolismo (ACE).

Funding: This project was funded by the Asociación Colombiana de Endocrinología, Diabetes y Metabolismo through an unrestricted educational grant from Adium. Methodological support was provided by Festina Lente, contracted by the Asociación Colombiana de Endocrinología, Diabetes y Metabolismo. The sponsors had no role in the design, conduct, or reporting of this review.

Conflicts of interest: The authors declare the following conflicts of interest. Natalia A. Rojas-Henao was contracted by Festina Lente for this work. Carlos E. Builes-Montaño has received consulting and speaking fees from Sanofi, Novo Nordisk, Novartis, Recordati Rare Diseases, Janssen, Chiesi, Abbott, and Boehringer Ingelheim, and is a shareholder of Festina Lente.

Acknowledgements: The authors acknowledge Festina Lente for providing methodological support, medical writing, and editorial assistance. The authors retain full responsibility for the content and conclusions of this review.

Authors' contributions: All authors contributed to the conception and design of the review. Carlos E. Builes-Montaño developed the search strategy. Natalia A. Rojas-Henao led the data extraction process and developed the computational tools and programming workflows used to support the screening, extraction, and analysis stages of the review. Natalia A. Rojas-Henao and Carlos E. Builes-Montaño provided methodological oversight. All authors contributed to the refinement of the manuscript and approved the final version of the manuscript.

Data availability statement: No primary data was collected in the development of this review. All data sources are publicly available literature.

Artificial intelligence disclosure: Artificial intelligence–assisted tools were used to support selected methodological and editorial tasks in this review. A custom Python workflow implemented in Google

Colaboratory used through the Google Generative Language application programming interface to generate provisional relevance summaries during title and abstract screening, assist with full-text eligibility assessment, and extract prespecified information into a structured data schema. OpenAI models were also used to support programming, data organization, synthesis development, figure-planning instructions, language editing, and manuscript drafting. Grammarly was used for language and style revision.

All artificial intelligence outputs were handled within a human-in-the-loop framework. Reviewers examined the original records and full-text articles and retained responsibility for all eligibility decisions, extracted values, study classifications, direction-of-effect judgments, critical appraisal assessments, and interpretations. No record was excluded, no data item was accepted as final, and no methodological judgment was assigned solely on the basis of an automated output. Machine-generated information was verified against the source publications, and discrepancies or ambiguous findings were corrected or adjudicated by the review team.

All descriptive analyses, tables, figures, and narrative conclusions were based on the final human-reviewed dataset. The authors critically reviewed and approved the complete manuscript and assume full responsibility for its accuracy, integrity, and conclusions.

References

1. Bjornsdottir S, Ing S, Mitchell DM, Sikjaer T, Underbjerg L, Hassan-Smith Z et al (2022) Epidemiology and Financial Burden of Adult Chronic Hypoparathyroidism. J bone mineral research: official J Am Soc Bone Mineral Res 37(12):2602–2614
2. Gosmanova EO, Houillier P, Rejnmark L, Marelli C, Bilezikian JP (2021) Renal complications in patients with chronic hypoparathyroidism on conventional therapy: a systematic literature review: Renal disease in chronic hypoparathyroidism. Rev Endocr Metab Disord 22(2):297–316
3. Büttner M, Singer S, Taylor K (2024) Quality of life in patients with hypoparathyroidism receiving standard treatment: an updated systematic review. Endocrine 85(1):80–90
4. Martin S, Chen K, Harris N, Vera-Llonch M, Krasner A (2019) Development of a Patient-Reported Outcome Measure for Chronic Hypoparathyroidism. Adv Ther 36(8):1999–2009
5. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Clinical Res ed) 372:n71
6. Iqbal K, Dass N, Gip C, Vila J, Rylands AJ, Marelli C (2019) Defining the Characteristics of Chronic Hypoparathyroidism Not Adequately Controlled on Conventional Therapy: Consensus Findings of Three European Delphi Panels. Adv Ther 36(11):3007–3016
7. Burillo E, Vila J, Oguiza A (2021) Characteristics of patients with chronic hypoparathyroidism not adequately controlled with conventional treatment and management patterns in Spain. Endocrinol Diabetes Nutr (Engl Ed) 68(5):304–311

8. Clarke BL (2022) Hypoparathyroidism: update of guidelines from the 2022 International Task Force. Arch Endocrinol Metab 66(5):604–610
9. Khan AA, Guyatt G, Ali DS, Bilezikian JP, Collins MT, Dandurand K et al (2022) Management of Hypoparathyroidism. J bone mineral research: official J Am Soc Bone Mineral Res 37(12):2663–2677
10. Khan AA, Bilezikian JP, Brandi ML, Clarke BL, Gittoes NJ, Pasieka JL et al (2022) Evaluation and Management of Hypoparathyroidism Summary Statement and Guidelines from the Second International Workshop. Journal of bone and mineral research: the official journal of the American Society for Bone and Mineral Research. ;37(12):2568–85
11. Bollerslev J, Buch O, Cardoso LM, Gittoes N, Houillier P, van Hulsteijn L et al (2025) Revised European Society of Endocrinology Clinical Practice Guideline: Treatment of Chronic Hypoparathyroidism in Adults. Eur J Endocrinol/Eur Fed Endocr Soc 193(5):G83–G112
12. Kassi E, Adamidou F, Yavropoulou MP, Anastasilakis AD, Makras P, Vryonidou A et al (2025) Diagnosis and management of hypoparathyroidism: recommendations of the working group of the Bone Section of the Hellenic Endocrine Society. Horm (Athens) 24(4):913–922
13. Makras P, Yavropoulou MP, Kassi E, Anastasilakis AD, Vryonidou A, Tournis S (2020) Management of parathyroid disorders: recommendations of the working group of the Bone Section of the Hellenic Endocrine Society. Horm (Athens) 19(4):581–591
14. Bollerslev J, Rejnmark L, Marcocci C, Shoback DM, Sitges-Serra A, van Biesen W et al (2015) European Society of Endocrinology Clinical Guideline: Treatment of chronic hypoparathyroidism in adults. Eur J Endocrinol/Eur Fed Endocr Soc 173(2):G1–20
15. Bilezikian JP, Brandi ML, Cusano NE, Mannstadt M, Rejnmark L, Rizzoli R et al (2016) Management of Hypoparathyroidism: Present and Future. J Clin Endocrinol Metab 101(6):2313–2324
16. Bollerslev J, Rejnmark L, Zahn A, Heck A, Appelman-Dijkstra NM, Cardoso L et al (2022) European Expert Consensus on Practical Management of Specific Aspects of Parathyroid Disorders in Adults and in Pregnancy: Recommendations of the ESE Educational Program of Parathyroid Disorders. Eur J Endocrinol/Eur Fed Endocr Soc 186(2):R33–r63
17. Maeda SS, Moreira CA, Borba VZC, Bandeira F, Farias MLF, Borges JLC et al (2018) Diagnosis and treatment of hypoparathyroidism: a position statement from the Brazilian Society of Endocrinology and Metabolism. Arch Endocrinol Metab 62(1):106–124
18. Mannstadt M, Clarke BL, Vokes T, Brandi ML, Ranganath L, Fraser WD et al (2013) Efficacy and safety of recombinant human parathyroid hormone (1–84) in hypoparathyroidism (REPLACE): a double-blind, placebo-controlled, randomised, phase 3 study. lancet Diabetes Endocrinol 1(4):275–283
19. Clarke BL, Vokes TJ, Bilezikian JP, Shoback DM, Lagast H, Mannstadt M (2017) Effects of parathyroid hormone rhPTH(1–84) on phosphate homeostasis and vitamin D metabolism in hypoparathyroidism: REPLACE phase 3 study. Endocrine 55(1):273–282

20. Mannstadt M, Clarke BL, Bilezikian JP, Bone H, Denham D, Levine MA et al (2019) Safety and Efficacy of 5 Years of Treatment With Recombinant Human Parathyroid Hormone in Adults With Hypoparathyroidism. J Clin Endocrinol Metab 104(11):5136–5147
21. Watts NB, Bilezikian JP, Bone HG, Clarke BL, Denham D, Levine MA et al (2023) Long-Term Safety and Efficacy of Recombinant Human Parathyroid Hormone (1–84) in Adults With Chronic Hypoparathyroidism. J Endocr Soc 7(5):bvad043
22. Khan AA, Rejnmark L, Rubin M, Schwarz P, Vokes T, Clarke B et al (2022) PaTH Forward: A Randomized, Double-Blind, Placebo-Controlled Phase 2 Trial of TransCon PTH in Adult Hypoparathyroidism. J Clin Endocrinol Metab 107(1):e372–e85
23. Khan AA, Rubin MR, Schwarz P, Vokes T, Shoback DM, Gagnon C et al (2023) Efficacy and Safety of Parathyroid Hormone Replacement With TransCon PTH in Hypoparathyroidism: 26-Week Results From the Phase 3 PaTHway Trial. J bone mineral research: official J Am Soc Bone Mineral Res 38(1):14–25
24. Clarke BL, Khan AA, Rubin MR, Schwarz P, Vokes T, Shoback DM et al (2025) Efficacy and Safety of TransCon PTH in Adults With Hypoparathyroidism: 52-Week Results From the Phase 3 PaTHway Trial. J Clin Endocrinol Metab 110(4):951–960
25. Tay YD, Tabacco G, Cusano NE, Williams J, Omeragic B, Majeed R et al (2019) Therapy of Hypoparathyroidism With rhPTH(1–84): A Prospective, 8-Year Investigation of Efficacy and Safety. J Clin Endocrinol Metab 104(11):5601–5610
26. Agarwal S, McMahon DJ, Chen J, Brossfield AV, Fernando J, Bilezikian JP et al (2023) The Clinical and Skeletal Effects of Long-Term Therapy of Hypoparathyroidism With rhPTH(1–84). J bone mineral research: official J Am Soc Bone Mineral Res 38(4):480–492
27. Rejnmark L, Gosmanova EO, Khan AA, Makita N, Imanishi Y, Takeuchi Y et al (2024) Palopegteriparatide Treatment Improves Renal Function in Adults with Chronic Hypoparathyroidism: 1-Year Results from the Phase 3 PaTHway Trial. Adv Ther 41(6):2500–2518
28. Rejnmark L, Gosmanova EO, Khan AA, Sprague S, Shoback DM, Kohlmeier L et al (2026) Sustained improvement in renal function with palopegteriparatide in adults with chronic hypoparathyroidism: 2-year results from the phase 3 PaTHway-trial. J bone mineral research: official J Am Soc Bone Mineral Res 41(9):922–933
29. Chen KS, Gosmanova EO, Curhan GC, Ketteler M, Rubin M, Swallow E et al (2020) Five-year Estimated Glomerular Filtration Rate in Patients With Hypoparathyroidism Treated With and Without rhPTH(1–84). J Clin Endocrinol Metab 105(10):e3557–e3565
30. Ayodele O, Rejnmark L, Mu F, Lax A, Berman R, Swallow E et al (2022) Five-Year Estimated Glomerular Filtration Rate in Adults with Chronic Hypoparathyroidism Treated with rhPTH(1–84): A Retrospective Cohort Study. Adv Ther 39(11):5013–5024
31. Rejnmark L, Ayodele O, Lax A, Mu F, Swallow E, Gosmanova EO (2023) The risk of chronic kidney disease development in adult patients with chronic hypoparathyroidism treated with rhPTH(1–84): A retrospective cohort study. Clin Endocrinol 98(4):496–504

32. Ayodele O, Mu F, Berman R, Swallow E, Rejnmark L, Gosmanova EO et al (2022) Lower Risk of Cardiovascular Events in Adult Patients with Chronic Hypoparathyroidism Treated with rhPTH(1–84): A Retrospective Cohort Study. Adv Ther 39(8):3845–3856
33. Sikjaer T, Rolighed L, Hess A, Fuglsang-Frederiksen A, Mosekilde L, Rejnmark L (2014) Effects of PTH(1–84) therapy on muscle function and quality of life in hypoparathyroidism: results from a randomized controlled trial. Osteoporos Int 25(6):1717–1726
34. Vokes TJ, Mannstadt M, Levine MA, Clarke BL, Lakatos P, Chen K et al (2018) Recombinant Human Parathyroid Hormone Effect on Health-Related Quality of Life in Adults With Chronic Hypoparathyroidism. J Clin Endocrinol Metab 103(2):722–731
35. Brandi ML, Vokes T, Appelman-Dijkstra NM, Ayodele O, Decallonne B, de Jongh R et al (2025) rhPTH(1–84) for hypoparathyroidism: a randomized study of patient-reported outcomes. Eur J Endocrinol/Eur Fed Endocr Soc 193(2):310–319
36. Rubin M, Palermo A, Vokes T, Khan AA, Schwarz P, Cetani F et al (2026) Patient-reported outcomes in palopegteriparatide-treated adults with hypoparathyroidism: PaTH Forward trial extension. J Clin Endocrinol Metab 111(4):945–952
37. Rubin MR, Clarke BL, Hofbauer LC, Khan A, Schwarz P, Vokes T et al (2026) Palopegteriparatide for Adults with Chronic Hypoparathyroidism: Skeletal Dynamics Through 3 year of the Phase 2 paTH Forward Trial. J bone mineral research: official J Am Soc Bone Mineral Res
38. Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M et al (2016) GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ (Clinical Res ed) 353:i2089
39. Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M et al GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ (Clinical research ed). 2016;353:i2016
40. Gattrell WT, Logullo P, van Zuuren EJ, Price A, Hughes EL, Blazey P et al (2024) ACCORD (ACcurate COnsensus Reporting Document): A reporting guideline for consensus methods in biomedicine developed via a modified Delphi. PLoS Med 21(1):e1004326
41. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J et al (1998) Consensus development methods, and their use in clinical guideline development. Health Technol Assess 2(3):i–iv

Figures

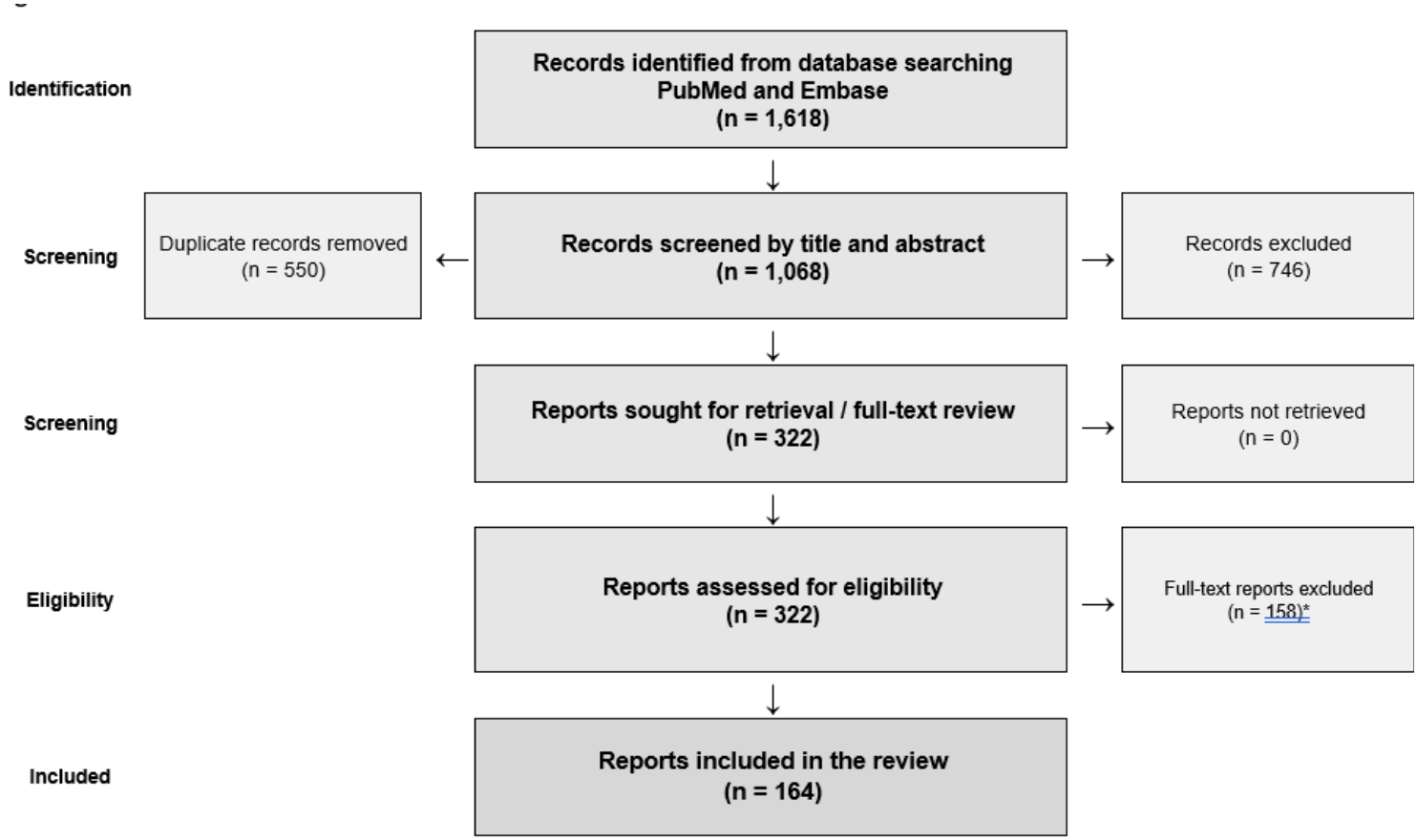


Figure 1

Figure 1. PRISMA 2020 flow diagram of study identification, screening, eligibility assessment, and inclusion.

Figure note. The search identified 1,618 records, of which 550 duplicates were removed. A total of 1,068 records underwent title and abstract screening; 746 were excluded and 322 reports proceeded to full-text assessment. All 322 reports sought for retrieval were obtained. Following full-text assessment, 157 reports were excluded and 165 were included in the review. The 165 included reports should not be interpreted as 165 unique studies because some represent companion publications, extensions, or secondary analyses of the same study or research program. Reasons for 22 of the 157 full-text exclusions remain to be prospectively coded before presentation of a complete exclusion-reason breakdown.

***** Full-text exclusion reasons were assigned to all 158 reports. This total included 22 reports excluded because no eligible comparison group was available for the comparative treatment-effect question. Absence of a comparator was not a general exclusion criterion for descriptive evidence. The complete distribution of full-text exclusion reasons is provided in Supplementary Table S11.

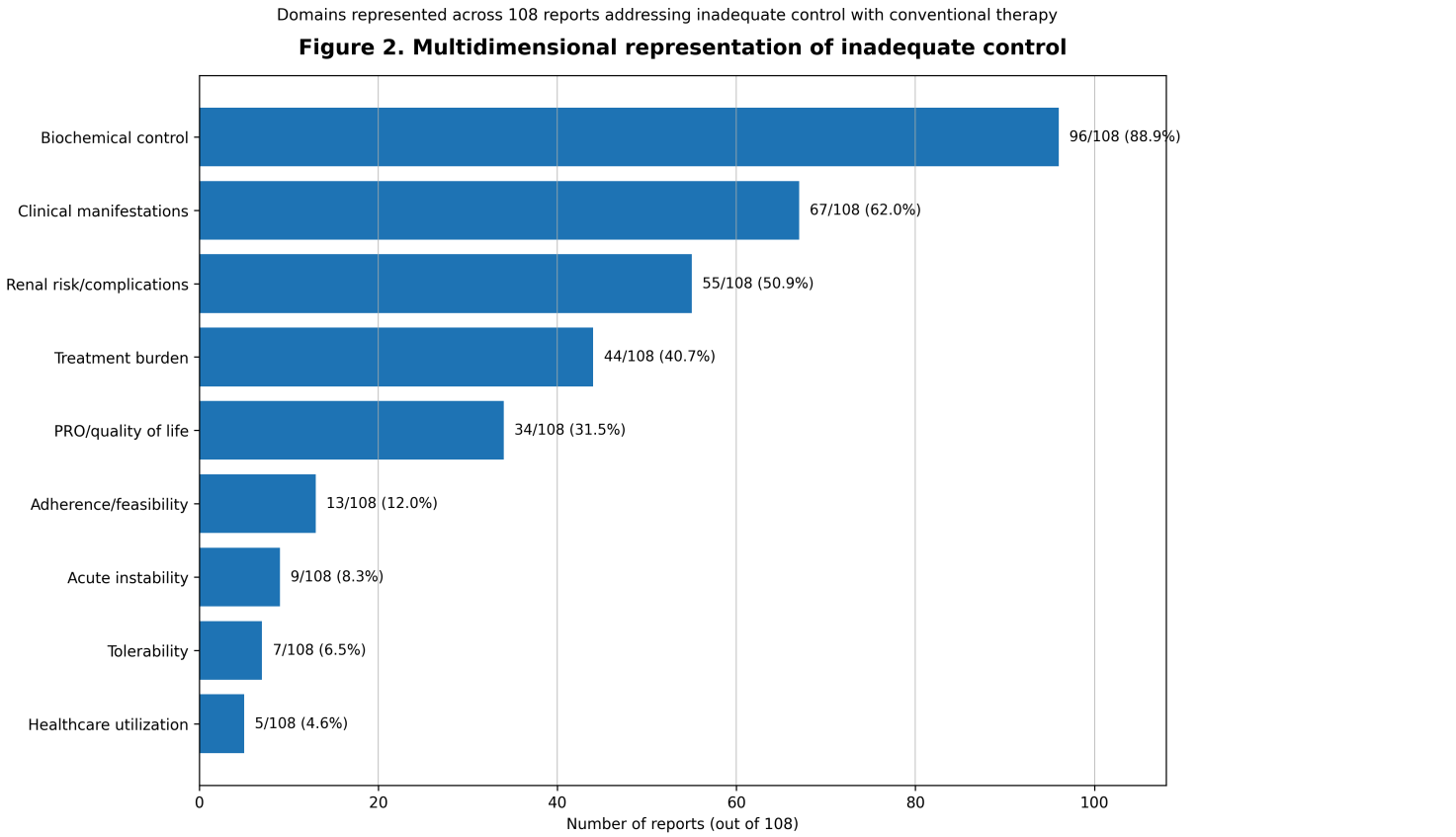


Figure 2

Figure 2. Multidimensional framework of inadequate control with conventional therapy in chronic hypoparathyroidism.

Figure note. “Inadequate control with conventional therapy” should be positioned centrally, surrounded by five recurrent core domains: biochemical control (96/108), clinical manifestations (67/108), renal risk/complications (55/108), treatment burden (44/108), and patient-reported outcomes/quality of life (34/108). Complementary domains should include adherence/feasibility (13/108), acute instability (9/108), tolerability (7/108), and healthcare utilization (5/108). Domain size may be visually proportional to reporting frequency, but

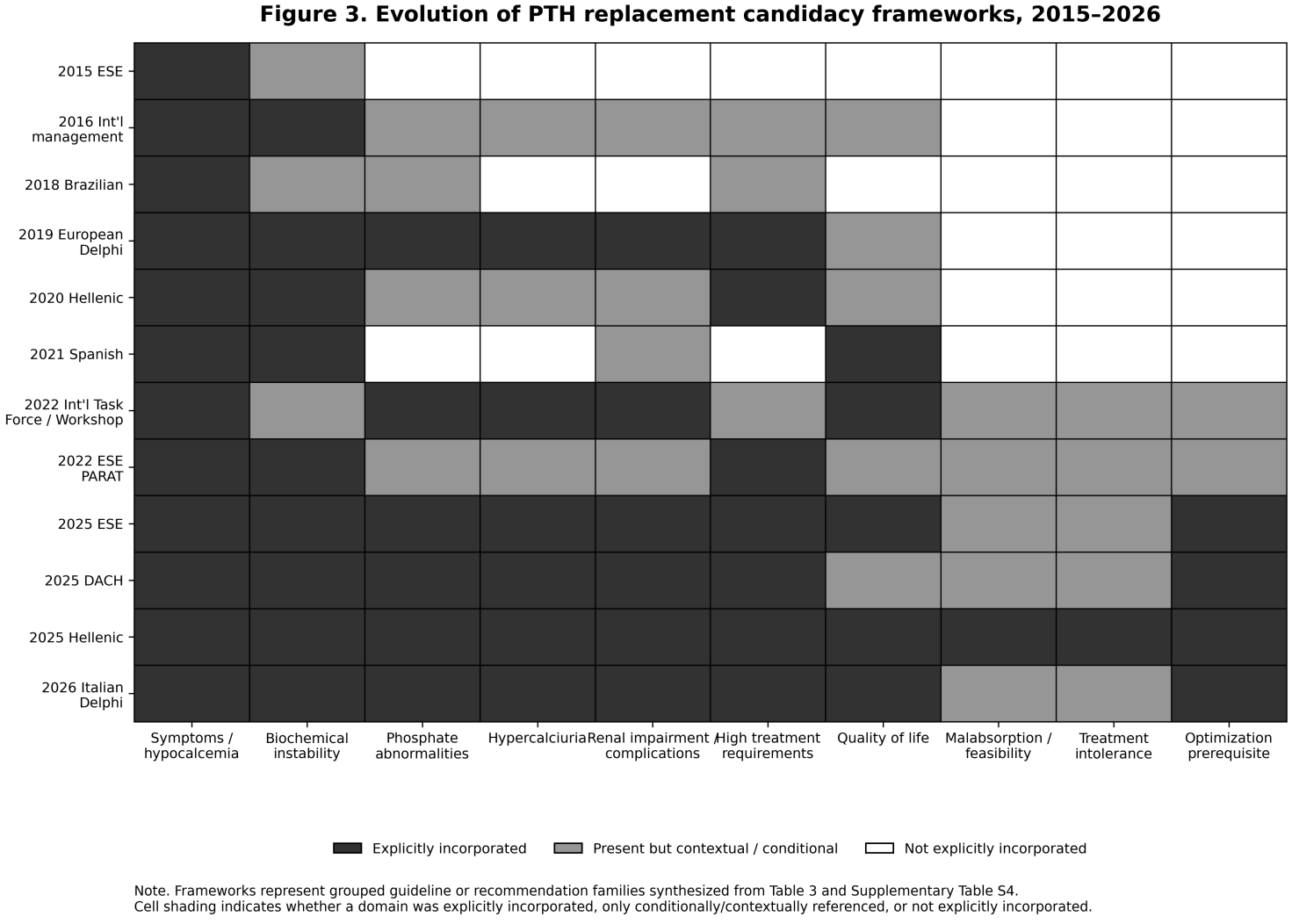


Figure 3

Figure 3. Evolution of domains incorporated into clinical candidacy frameworks for PTH replacement therapy in chronic hypoparathyroidism, 2015–2026.

Figure note. The figure should display major guidance or consensus frameworks chronologically and indicate whether each incorporated the following domains into consideration of PTH replacement: persistent hypocalcemia or symptoms, biochemical instability, phosphate abnormalities, hypercalciuria, renal impairment or renal complications, high conventional-treatment requirements, impaired quality of life, malabsorption/feasibility, and treatment intolerance. Related publications from the same guideline framework should be represented once. The figure is intended to illustrate the progressive broadening and heterogeneity of candidacy constructs and should not be interpreted as a hierarchy of evidence, recommendation strength, or preferred clinical algorithm.

A timeline/heat-map hybrid would work particularly well: rows for domains, columns for guideline years, with a marker indicating whether that domain was explicitly incorporated.

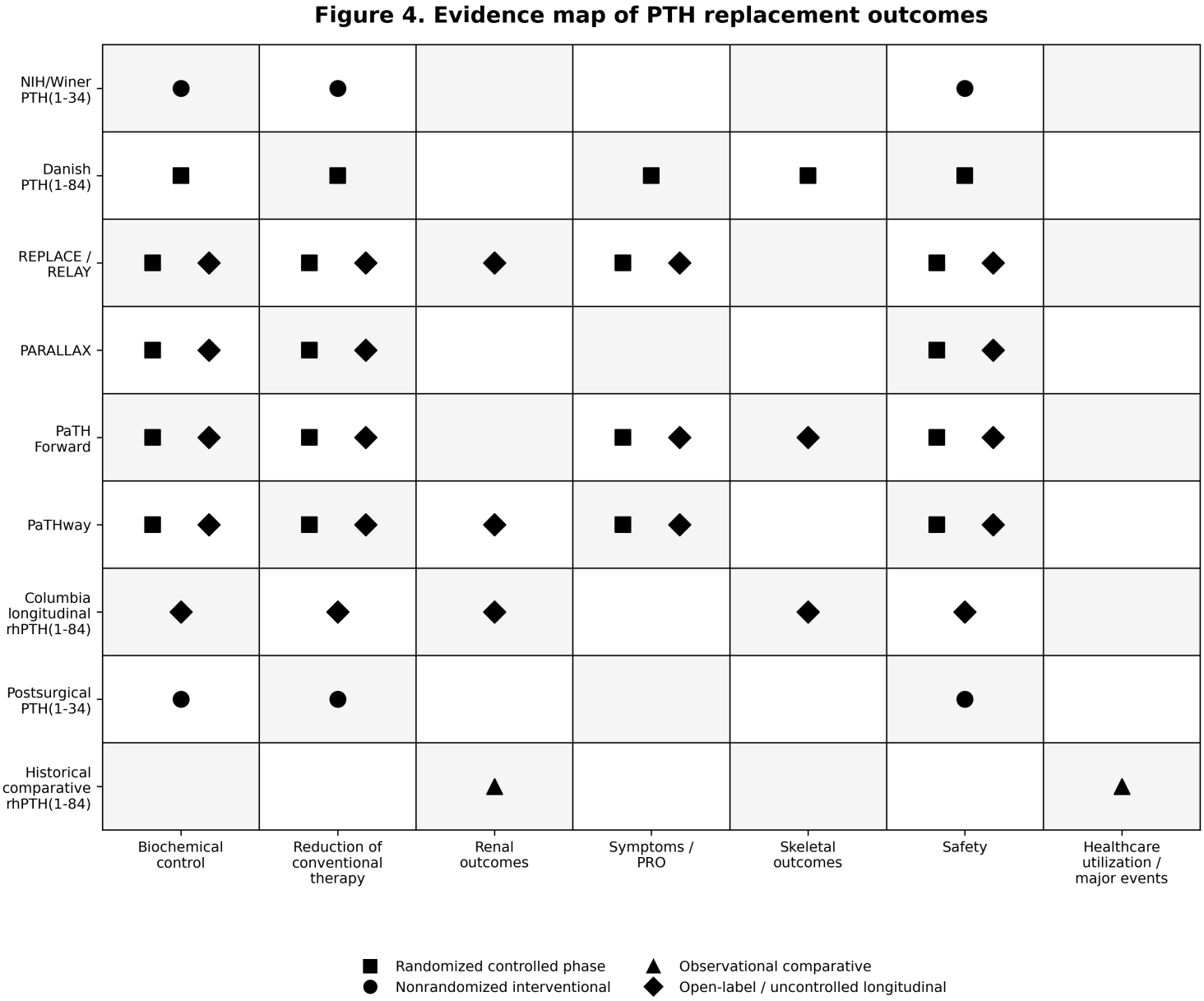
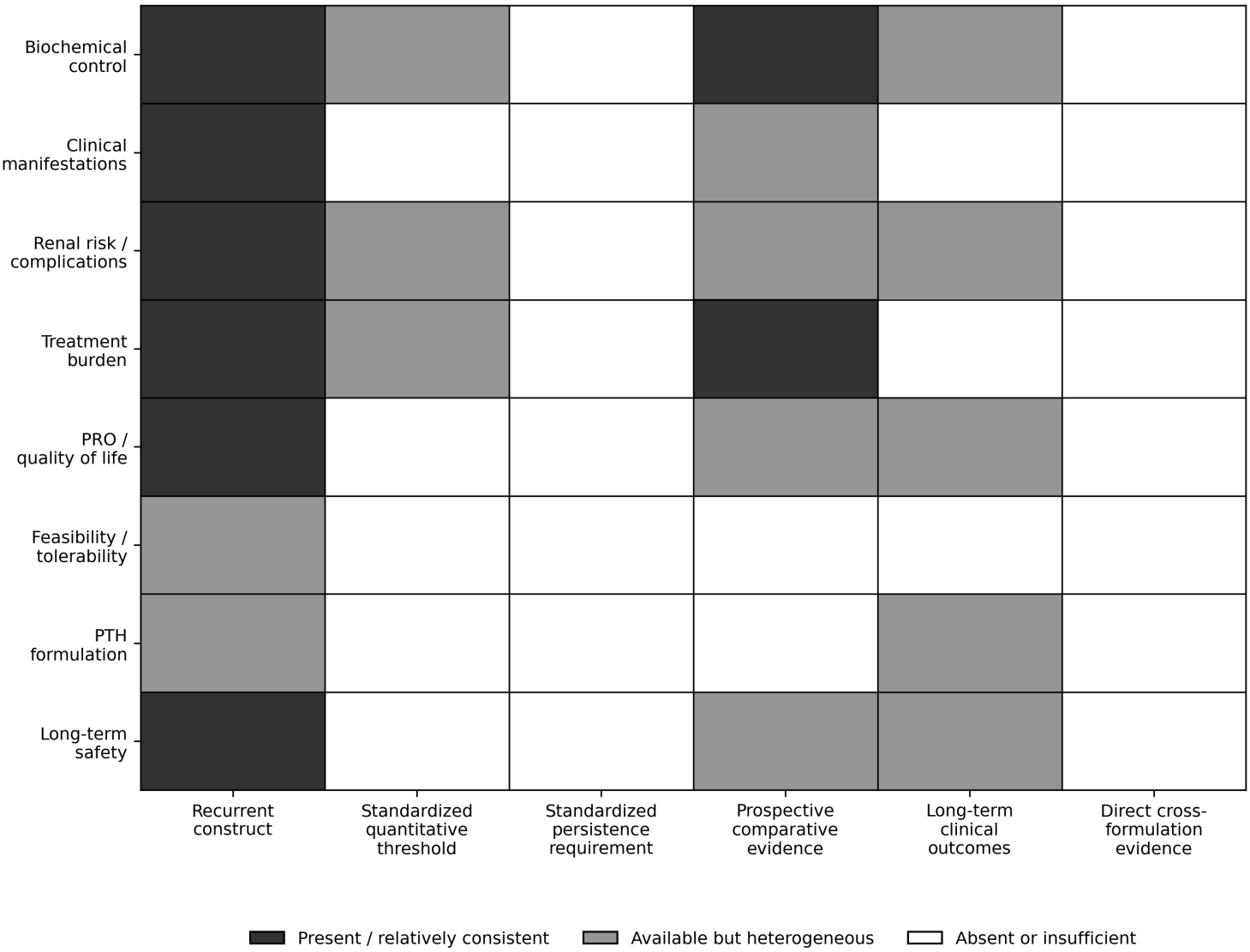


Figure 4

Figure 4. Evidence map of PTH replacement therapy outcomes by formulation and outcome domain.

Figure note. Rows should represent PTH formulation or major study program and columns should represent biochemical control, reduction of conventional therapy, renal outcomes, symptoms/patient-reported outcomes, skeletal outcomes, safety, and healthcare utilization or major clinical events. Cell symbols should indicate study design—randomized controlled, nonrandomized interventional, observational comparative, or uncontrolled longitudinal—rather than direction or magnitude of effect. Companion publications from the same study program should be linked within a single row or visual cluster. The figure should not be interpreted as a certainty-of-evidence rating.

Figure 5. Evidence-gap map



Note. Cell shading represents availability and standardization of evidence, not methodological quality, certainty of evidence, recommendation strength, or clinical importance.

Figure 5

Figure 5. Evidence-gap map for operationalizing inadequate control and PTH replacement candidacy.

Figure note. Rows should represent the principal decision domains: biochemical control, clinical manifestations, renal risk/complications, treatment burden, patient-reported outcomes, feasibility/tolerability, PTH formulation, and long-term safety. Columns should represent whether the literature provides: (1) a recurrent construct, (2) a standardized quantitative threshold, (3) a standardized persistence requirement, (4) prospective comparative evidence, (5) long-term clinical outcomes, and (6) direct cross-formulation evidence. Symbols should indicate presence, absence, or heterogeneity of the relevant evidence feature; they should not be interpreted as ratings of evidence certainty or clinical importance.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- HypoparathyroidismSupplementaryMaterialcomplete.docx