Preprints.org

Article

Not peer-reviewed version

Rising Use of Unapproved BPC-157 (“Wolverine Peptide”) and Associated Patient-Reported Outcomes Curated from Clinical Notes

AJ Venkatakrishnan , Adhikaar Marwaha , Christopher J. Gregg , Gowtham Varma , Matthew Hurchik , Karthik Murugadoss , Santhosh Shivabasappa , Venky Soundararajan *

Posted Date: 7 September 2026

doi: 10.20944/preprints202609.0501.v1

Keywords: BPC-157; peptides; unapproved drugs; gray-market therapeutics

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Rising Use of Unapproved BPC-157 ("Wolverine Peptide") and Associated Patient-Reported Outcomes Curated from Clinical Notes

AJ Venkatakrishnan [1,2], Adhikaar Marwaha [3], Christopher J. Gregg [1], Gowtham Varma [3], Matthew Hurchik [1], Karthik Murugadoss [1,2], Santhosh Shivabasappa [3] and Venky Soundararajan [1,2,3,*]

[1] Nference, 1 Main Street, East Arcade building, Cambridge, Massachusetts, USA

[2] Metabolism Agentic Intelligence Atlas (MAIA), Cambridge, Massachusetts, USA

[3] Nference Labs, 4th Floor, Indiqube, Golf View Campus Tower-2, Bengaluru, India

* Correspondence: venky@nference.net

Abstract

Body Protection Compound-157 (BPC-157) is an unapproved synthetic peptide marketed directly to consumers despite limited human evidence on benefits and risks. We characterized the use of BPC-157 and patient-reported outcomes from 2020 through 2026 using Large Language Model (LLM) curation of de-identified clinical notes from a U.S. federated network with physician validation of each LLM-curation exercise. Of 1,536 patients with a note mentioning BPC-157, 1,039 (67.6%) had documented use, and quarterly newly confirmed users increased 33-fold between 2020 and 2026. Among 972 BPC-157 users with recorded race, 95.9% were White, compared with 78% of the background care population, representing a 1.23-fold enrichment (chi-square test across racial categories, $P < 0.001$). Among patients with newly documented BPC-157 use, the male proportion increased from 57% in 2020 to 70% in 2026 (Cochran–Armitage trend test across years, $P < 0.001$), while mean age decreased from 55 years in 2020 to 49 years in 2026. Among 205 patients with a documented consumer source (19.7% of confirmed users), BPC-157 was most commonly obtained from compounding pharmacies (36%) or gray-market peptide vendors (35%). Other therapeutic agents were co-used with BPC-157 by 525 of 1,039 users (50.5%), most commonly testosterone (17.0%), NSAIDs (14.7%), TB-500 (12.9%), and corticosteroids (11.7%), while physical therapy (15.3%), surgery (12.6%), exercise counseling (8.3%), and diet counseling (6.3%) were the documented non-drug co-interventions. Physician-adjudicated extraction accuracy was 96.0% for BPC-157 co-therapies and 74.4% for non-drug interventions, which were treated as exploratory. Reasons for use were documented for 644 patients and included pain (33% of all 1,039 users), gastrointestinal conditions including reflux, inflammatory bowel disease or gastritis (14%), prior injury (11%), and post-surgical healing (10%), physician adjudicated LLM accuracy of 83.9%. Among 80 users with ascertainable start timing, BPC-157 use preceded first clinician documentation by a mean of 60 days. Reported use averaged 2.2 months among 130 patients (12.5%) with an explicitly documented BPC-157 use duration, with 96.7% extraction accuracy. The specialties most often first documenting use were family medicine, internal medicine, gastroenterology, and orthopaedic surgery. Symptomatic improvement of pain, wound or tissue healing and functional return was documented in 79% of 354 BPC-157 users with a directional response, and symptomatic worsening in 5% of the users (LLM-extraction accuracy 78.4%). Response direction was unavailable for 685 of 1,039 users (65.9%). These reports cannot establish efficacy or distinguish BPC-157 effects from peptide stacking, other co-treatments, non-drug interventions, placebo effects or selection, reporting and confirmation biases. LLM-curated adverse events during BPC-157 use were infrequently noted and likely subject to under-reporting, with documented evidence of neuropsychiatric (1.4%), injection-site hypersensitivity (1.4%), and gastrointestinal (1.3%) events, with physician-adjudicated accuracy of 93.9% for drug-attributed adverse-events. With BPC-157 use rising rapidly despite minimal prospective studies and

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

limited evidence-backed consumer awareness of its risk–benefit profile, this study highlights the need for facilitating structured EHR capture of broader gray-market peptide use and a continuous assessment of guidelines to better support healthcare practitioners.

Keywords: BPC-157; peptides; unapproved drugs; gray-market therapeutics

1. Introduction

BPC-157 (body-protection compound-157) is a synthetic pentadecapeptide marketed directly to consumers for musculoskeletal healing, soft-tissue repair, and gastrointestinal complaints, despite the absence of large-scale human evidence and any regulatory approval for these indications. Its scientific rationale is derived largely from preclinical work: experimental studies have reported enhanced tendon fibroblast migration and survival, improved ligament healing in rats, and pro-angiogenic VEGFR2–Akt–eNOS signaling [1–4]. Public and commercial interest in BPC-157 and related peptides has grown rapidly, amplified by a recent U.S. Food and Drug Administration advisory-committee discussion of the compounding status of several peptides [5].

The clinical evidence base for BPC-157 remains strikingly meager relative to public interest. Published reports include a retrospective knee-pain series, a 12-patient interstitial-cystitis pilot and a two-participant intravenous safety pilot, with some earlier placebo-controlled enema studies being largely under-powered [1,6–8]. The US Food and Drug Administration (FDA) has not approved any BPC-157 product and has emphasized limited clinical safety information, physicochemical uncertainty, peptide-related impurities, aggregation and potential immunogenicity as sources of concern for prospective users [5]. BPC-157 is also named in the World Anti-Doping Agency's S0 class of non-approved substances [9].

The combination of rapid consumer adoption, uncertain product quality and lack of conventional clinical regulation creates a measurement problem for real world study. Because BPC-157 is unapproved and can be obtained outside conventional prescribing pathways, its use is expected to be absent from structured medical records. Clinical notes may therefore provide the only Electronic Health Record (EHR) evidence confirming use and describing product source, reason for use, perceived benefit and possible adverse effects. EHR phenotyping of BPC-157 use therefore requires extraction of temporality, assertion, attribution and context from free text of clinical notes [10,11]. Large language models (LLMs) can perform document-level clinical summarization and adverse-event extraction with high accuracy when validated against expert review. However, model error and documentation bias remain material, necessitating expert physician adjudication [11–13].

Here, we used LLM-based curation as the primary approach to ascertain BPC-157 exposure from clinical notes and characterize the real-world contexts in which it is used. We focused on three questions of broad clinical and public health relevance: how rapidly documented use is increasing, where patients report obtaining BPC-157 and the reasons they use it, and what symptomatic or functional outcomes are reported following exposure. Together, these analyses provide a real-world view of an increasingly visible but poorly characterized unapproved peptide whose use largely occurs outside conventional prescribing pathways.

2. Methods

2.1. Study Design and Data Source

We conducted a retrospective observational study of de-identified EHR data using the nference federated network. The study population comprised patients with any mention of BPC-157 in their clinical documentation through the second quarter of 2026. Because BPC-157 has no approved product, it has no National Drug Code and no structured prescription record; exposure ascertainment therefore was derived from clinical notes, and the complete clinical note history of each patient with a BPC-157 mention was retrieved for adjudication. No minimum post-index observation period was

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

required for cohort entry; follow-up was as-observed, and each response analysis was restricted to the subset of users with the relevant post-index documentation.

2.2. Exposure Ascertainment from Clinical Notes

A large language model (LLM) adjudicated each patient's full note text to determine whether the patient actually took BPC-157, distinguishing confirmed exposure from discussion without use, denial of use, and unclear documentation. Whole notes rather than keyword windows were supplied as context, because procurement, indication, and response details frequently appear in social-history, medication-reconciliation, and plan sections distant from the drug name. The auditable LLM-curation methodology, including its validation against structured records for incretin exposure ascertainment, has been described previously [11,14,15]; model configuration and the full text of every extraction prompt are provided in the Supplementary Materials. Each patient's index date was the earliest note confirming exposure. Where a note stated an interpretable time since starting BPC-157, the reported start was resolved to a calendar span.

2.3. Supply Route, Indication, and Treatment-Pattern Curation

Additional LLM passes over the confirmed users' notes classified the documented supply route (consumer channels: compounding pharmacy, peptide or gray-market vendor, wellness/medical-spa/weight clinic, online or telehealth vendor, over-the-counter retail, or friends or family, versus clinician prescription, clinical-trial supply, or other), the reason or indication for use (with the "no specific reason stated" category assigned if the note documented use without a stated reason), the injury site (canonicalized to a single musculoskeletal joint per patient, merging laterality and excluding non-injury visceral sites), the post-surgical procedure region,the route of administration, and concomitant interventions. Monotherapy versus stacking was defined by whether there was documented co-use of a fixed, pre-specified set of therapeutic agents (TB-500, CJC-1295, ipamorelin, growth hormone, insulin-like growth factor 1, testosterone, selective androgen-receptor modulators, anabolic steroids, platelet-rich plasma, corticosteroids, non-steroidal anti-inflammatory drugs, semaglutide, tirzepatide, and any other peptide); co-use of a named peptide was counted only where the same patient was independently confirmed as a user of that peptide by the identical note-based ascertainment, rather than inferred from an unverified mention; "monotherapy" denotes the absence of any documented co-use of these agents. The prompts associated with this curation are provided in the Supplementary Materials.

2.4. Symptomatic and Functional Response

For each confirmed user, the LLM classified the documented direction of symptomatic or functional response to BPC-157 (improved, no change, worse, or not stated), the reporter of that response (patient, clinician, or both), the specific symptom and functional-improvement types, and the affected quality-of-life and functional domains. The assessable-response denominator comprised users with a documented response direction for each of these domains.

2.5. Physician Validation

Accuracy of the LLM extractions was assessed by manual review at one participating site. For each core extraction, items were sampled at random, stratified by the value the model had assigned, so that class-specific accuracy could be estimated. A clinician reviewer, presented with the relevant note excerpt alongside the extracted value, recorded each extraction as correct, incorrect, or indeterminate; agreement was computed within each predicted class and combined by class prevalence, with Wilson score confidence intervals [16] (Table 3). Agreement was highest for reported duration of use (96.7%), documented symptomatic response (88.2%), exposure ascertainment (87.2%), and supply route (93.8%), and it was more moderate (approximately 65–80%) for injury site, musculoskeletal-indication subtype, and benefit type, which require finer free-text categorization.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

Extractions with physician-validated agreement below 80% are treated as exploratory and are not used to support mechanistic claims; the corresponding figures should be read descriptively.

2.6. Statistical Analysis

Proportions are reported with Wilson score 95% confidence intervals [16]. Mean age across initiation years was compared by one-way analysis of variance; categorical distributions (for example, race versus the source population) by the chi-square test; ordered strata (the documented-use duration–response gradient) by the Cochran–Armitage trend test [17,18]. Multiplicity was controlled by the Benjamini–Hochberg false-discovery-rate procedure [19]. within each prespecified comparison family, and q values are reported alongside P values ($q < 0.05$ considered significant and shown in bold in figures and tables). All estimates are pooled across the federated network from additive sufficient statistics; because distinct patients do not span sites, pooled counts equal the sum of site counts. Any cell with a patient count below 11 is reported as "<11" with derived percentages suppressed, per the privacy convention adopted in this study. Endpoints, comparison families, and small-cell suppression rules were pre-specified before the outcome analyses were run, and the study is reported in accordance with the STROBE and RECORD reporting guidelines [20,21].

2.7. Institutional Review Board Statement, Informed Consent Statement, De-Identification and HIPAA Compliance Certification

Prior to analysis, all EHR data were de-identified under an expert determination consistent with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule (45 CFR §164.514(b)(1)). The de-identification methodology employed a multi-layered transformation approach to both structured and unstructured data fields. In structured data, direct identifiers including patient names and precise geographic locations were excluded entirely, while indirect identifiers underwent specific transformations: patient identifiers, medical record numbers, and accession numbers were replaced with one-way cryptographic hashes using confidential salts to preserve linkage across patient encounters; all dates were shifted backward by patient-specific random offsets (1 to 31 days) to preserve temporal relationships while obscuring exact event timing; ZIP codes were truncated to two-digit state-level resolution; and continuous variables including age, height, weight, and body mass index were thresholded to prevent identification of extreme values (for example, ages 89 years or older transformed to "89+" and body mass index over 40 transformed to "40+"). In clinical text, an ensemble de-identification system combining attention-based deep learning with rule-based methods achieved an estimated >99% recall for personally identifiable information detection [22], with detected identifiers replaced by plausible fictional surrogates. Institutional Review Board Statement and Informed Consent Statement are not applicable.

2.8. Data Harmonization

To address heterogeneity in EHR data, we harmonized clinical variables including medications, anthropometric measurements, and diagnoses to standardized concepts. For medications, we constructed a standardized drug-concept database combining the nference knowledge graph with RxNorm hierarchies to capture ingredient, brand, and dose-specific information; medication records were matched hierarchically, prioritizing RxNorm codes, then ingredient-level matching, and finally natural-language processing and pattern matching on free-text orders when structured codes were absent. For anthropometric measurements, we created a unified vocabulary from SNOMED and LOINC and matched measurement descriptions using standardized text matching with abbreviation expansion and synonym resolution, resolving ambiguous mappings with OpenAI GPT-4o [23]. using summary statistics as context followed by manual verification. For diagnoses, we developed a hierarchical disease-concept database from the knowledge graph and matched records by identifying the most specific common child concept in the hierarchy.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

2.8.1. Data Availability

This study analyzed de-identified electronic health record data. The data were extracted under an established privacy-preserving protocol and cannot be shared beyond what is reported here due to the parameters of the expert determination.

2.8.2. Code Availability

The analysis code is not publicly available. Please contact the corresponding author for details.

3. Results

Among 15.2 million patients in the federated data network with at least one clinical note between 2020 and 2026, 1,536 (0.01%) had at least one note that mentioned BPC-157 (Figure 1A). Of these patients, LLM-based curation ascertained 1,039 (67.6%) patients with confirmed administration ("analytic cohort"), whose demographic characteristics are summarized in Table 1. The average age was 49.6 years (SD: 14.6 years), and 59.5% of patients with a documented sex were male, versus 46.7% in the background population (Table 1). The documented race was white for 95.9% of patients, versus 78% among patients with a documented race in the background population (Figure 2C, Table 1). The average baseline body mass index (BMI) was 28.0 kg/m2 (SD: 5.4 kg/m2). Manual review of a sample of 50 notes mentioning BPC-157 showed the accuracy of the LLM to be 87% for the determination of confirmed administration (Table 3). Temporal analysis demonstrated a notable increase in confirmed BPC-157 patients in the past five years. The number of newly confirmed BPC-157 patients rose steeply over the study period (Figure 1B), increasing approximately 33-fold from 4 in the first quarter of 2020 to 134 in the first quarter of 2026. Interestingly, the demographic composition of patients who have reported BPC-157 has evolved over the study period, with the percentage of male patients increasing from 57% in 2020 to 70% in 2026 and the mean age of all patients decreasing from 55 years to 49 years (Figure 2A,B).

Table 1. Demographic and baseline clinical characteristics of confirmed BPC-157 patients (N = 1,039), with the background network population where comparable. Percentages are of the stated denominator; means are shown with SD. Background sex and race are shown among individuals with a documented value, matching the cohort denominators. Background age is the current age (2026 minus birth year) across persons with a recorded birth year, whereas the cohort age is at first documented exposure. Relative to the background population, confirmed BPC-157 patients showed 1.27-fold enrichment of male sex and 1.23-fold enrichment of White race.

Characteristic	Confirmed BPC-157 patients (N = 1,039)	Background network population
Age at first documented exposure, mean (SD), years	49.6 (14.6)	55.5 (27.5)
Sex — of 1,023 with a documented sex, n (%):		
Female	414 (40.5)	53.3%
Male	609 (59.5)	46.7%
Race — of 972 with a documented race, n (%):		
White	932 (95.9)	78.3%
Black or African American	16 (1.6)	16.6%
Other or multiple	24 (2.5)	5.0%

BPC-157 was commonly used as part of a broader treatment regimen rather than as an isolated intervention. Among 1,039 confirmed users, 525 (50.5%) had documented co-use of at least one other therapeutic agent, whereas 355 (34.2%) used BPC-157 alone without a documented concomitant therapeutic or lifestyle intervention (Figure 1C). Among patients not stacking BPC-157 with other

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

therapeutic agents, documented adjunctive management included physical therapy (6.4%, n=66), surgery (5.3%, n=55), exercise counseling (4.0%, n=42), and diet counseling (3.5%, n=36); these lifestyle or procedural interventions were not necessarily mutually exclusive (Figure 1C).

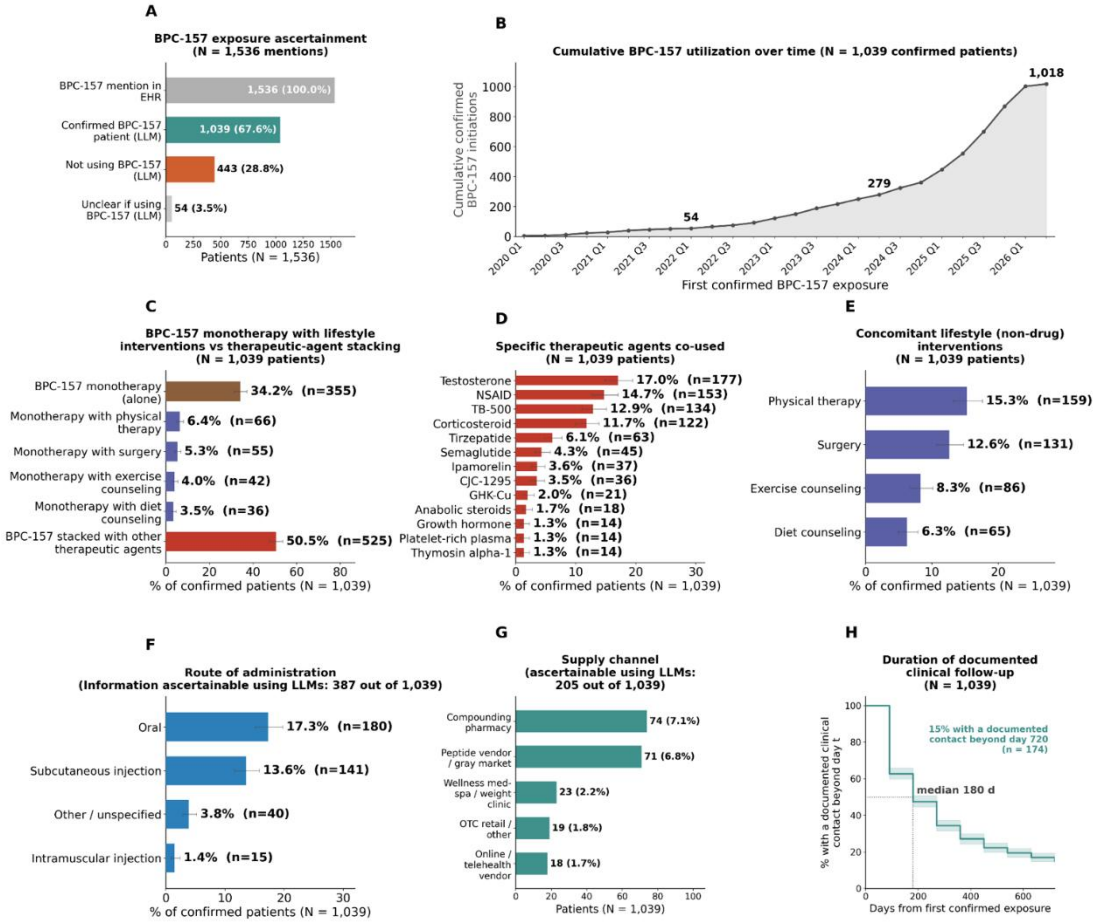


Figure 1. LLM-ascertained sourcing and patterns of BPC-157 use. Panel A summarises the large-language-model exposure adjudication, panels C–G are large-language-model-ascertained note curation, and panels B and H derive from dated clinical records. (A) Exposure-ascertainment status among patients with a BPC-157 mention (mention → note-confirmed patient → determined not to be using → unclear). (B) Cumulative documented BPC-157 use over time across the federated data network. (C) BPC-157 used as monotherapy versus stacked with at least one other therapeutic agent; monotherapy combined only with a lifestyle (non-drug) intervention is shown as separate bars. (D) Specific therapeutic agents co-used with BPC-157. (E) Concomitant non-drug (lifestyle) interventions across the full confirmed cohort (denominator: all confirmed patients) — distinct from panel C, whose lifestyle bars are limited to the BPC-157 monotherapy subgroup. (F) Documented route of administration. (G) Supply channel among patients with a documented consumer source. (H) Documented clinical follow-up — a Kaplan–Meier estimate of the time from first confirmed BPC-157 exposure to each patient's last documented clinical note of any type (any note in the electronic health record, not restricted to BPC-157 mentions), with a Greenwood 95% confidence band and the median annotated; the curve is shown to the horizon at which at least 15% of the cohort still has documentation. It therefore measures how long each patient remains observable in the record after first exposure (the follow-up window) and is a proxy for loss to documented follow-up — not BPC-157 persistence, active care, or survival. The absence of a later note reflects the end of available documentation in the network, not necessarily disengagement from care, death, or relocation. No minimum follow-up period was required for cohort entry, so follow-up is as-observed. For panels where a metric could not be ascertained for every patient (route, supply channel), percentages use the full confirmed cohort as the denominator and the ascertainable count is stated in the panel heading. Proportions carry Wilson 95% confidence intervals; counts below 11 are suppressed.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

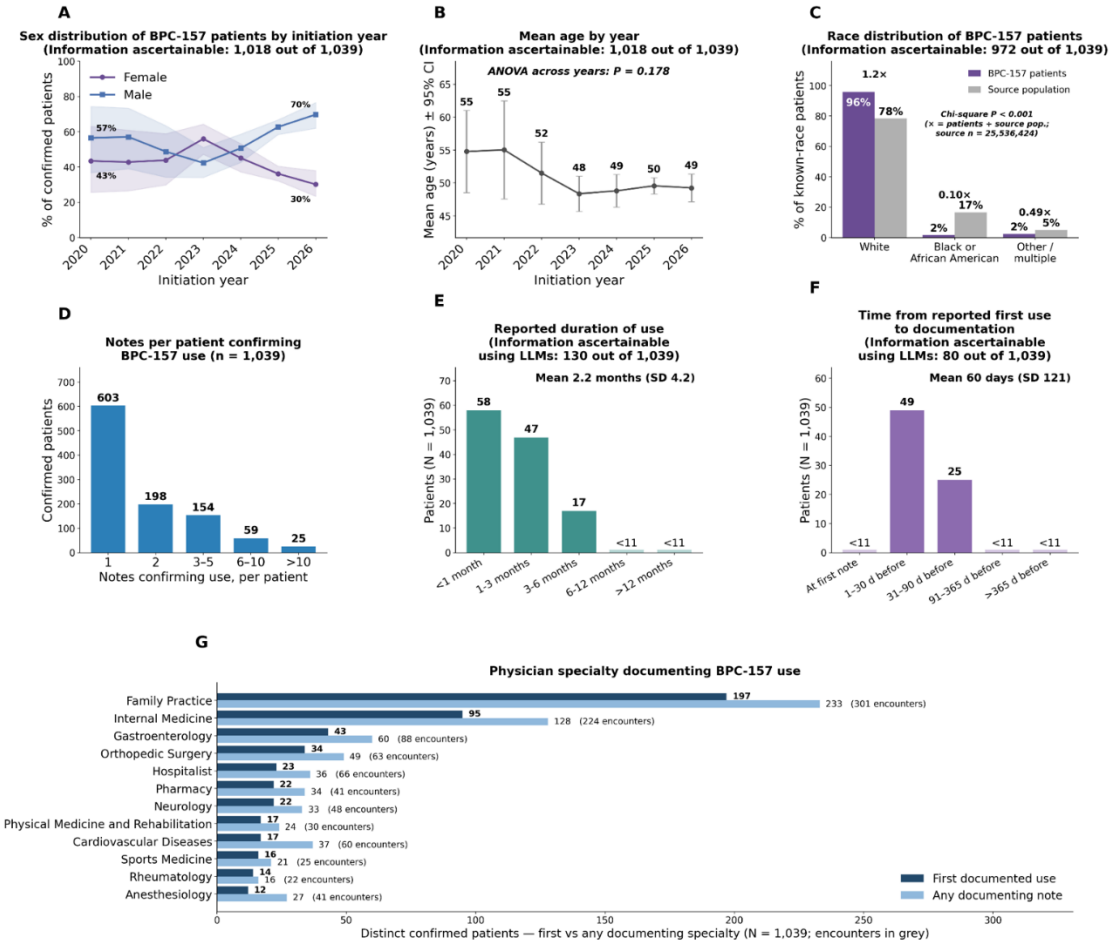


Figure 2. Demographics of BPC-157 patients and documentation of use (N = 1,039). (A) Documented sex by year of first confirmed exposure, with Wilson 95% confidence-interval ribbons. (B) Mean age at first confirmed exposure by year with 95% confidence intervals; the difference across years was tested by one-way analysis of variance. (C) Race and ethnicity of confirmed patients versus the source population of the same health systems (restricted to patients with a recorded race); the difference was tested by the chi-square test. (D) Distribution of the number of notes confirming use per patient. (E) LLM-ascertained reported duration of use, among patients for whom a duration was stated in the record (a reported duration such as "a few weeks," not a count of notes). (F) Time from the patient-reported start of use to the first clinical encounter that documented such use. (G) Documenting-clinician specialty: the specialty at the first documented BPC-157 use (each patient counted once, in the specialty that first surfaced the use) shown alongside any documenting specialty (a patient may be documented by more than one), with the per-specialty encounter count in grey. Family and internal medicine predominate but are expected to be over-represented as primary-care defaults, where patients are commonly seen before specialist routing, so the first-documented view is the less-confounded read. Counts below 11 are not shown explicitly.

The most frequently co-used therapeutic agent was testosterone, documented in 17.0% of BPC-157 users (n=177), followed by nonsteroidal anti-inflammatory drugs (NSAIDs; 14.7%, n=153), TB-500 (12.9%, n=134), and corticosteroids (11.7%, n=122) (Figure 1D). Metabolic therapies were also present, including tirzepatide (6.1%, n=63) and semaglutide (4.3%, n=45). Additional peptide or hormone-related agents included ipamorelin (3.6%, n=37), CJC-1295 (3.5%, n=36), GHK-Cu (2.0%, n=21), anabolic steroids (1.7%, n=18), growth hormone (1.3%, n=14), and thymosin alpha-1 (1.3%, n=14). Platelet-rich plasma was documented in 1.3% of users (n=14) (Figure 1D). Physician-adjudicated accuracy for co-therapy ascertainment was 96.0% (Table 3).

Beyond pharmacologic co-use, BPC-157 users frequently received conventional supportive or procedural care. Physical therapy was documented in 15.3% of patients (n=159), surgery in 12.6%

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

(n=131), exercise counseling in 8.3% (n=86), and diet counseling in 6.3% (n=65) (Figure 1E). Together, these findings indicate that BPC-157 exposure commonly occurred within multimodal treatment strategies incorporating other drugs, peptides, procedures, and rehabilitation, rather than as a standalone intervention.

The most common documented route of administration was oral (180 of 1,039, 17.3%), followed by subcutaneous injection (141 of 1,039, 13.6%) and intramuscular injection (15, 1.4%) (Figure 1F). The source through which BPC-157 was obtained was documented for 205 of the 1,039 (19.7%) patients in the analytic cohort. The most common sources included compounding pharmacies (74 of 205, 36%) and gray-market peptide vendors (71 of 205, 35%); less common sources included wellness spas or weight clinics (23, 11%), retail or other outlets (19, 9%), and online or telehealth vendors (18, 9%) (Figure 1G). Supply-channel ascertainment had a physician-adjudicated accuracy of 93.8% (Table 3). The route of administration could not be ascertained for the remaining 652 (63%) patients. Importantly, we were not able to reliably assess dose quantities or adherence patterns given the intrinsically patient-reported nature of the data for this non-approved compound (i.e., absence of structured records for prescription orders and fills). No minimum follow-up period was required for cohort entry; the median documented follow-up (first confirmed exposure to last documented contact) was 180 days, and 15% of patients had a documented clinical contact beyond two years (Figure 1H).

The distribution of the number of notes confirming BPC-157 use per patient is shown in Figure 2D, with 603 of 1,039 (58%) patients having only one such note, 19% having two notes, 15% having three to five notes, and few patients with six or more notes. Duration of use was explicitly documented for 130 of the 1,039 (12.5%) patients, with most of these patients reporting less than three months of use (physician-adjudicated accuracy 96.7%; Figure 2E, Table 3). In 80 patients for whom a specific start date was provided, reported BPC-157 initiation typically preceded the first documenting note by weeks to months, with an average of 60 days (physician-adjudicated accuracy 94.0%; Figure 2F, Table 3). BPC-157 use was most often first documented by family- and internal-medicine clinicians (Figure 2G). Documentation, use-pattern, and baseline injury-characterization details for the cohort — including supply route, use duration, symptomatic response, and injury type, severity, chronicity, prior treatments, and healthcare utilization — are summarized in Table 2.

Table 2. Documentation, use patterns, and documented outcomes among confirmed BPC-157 patients (N = 1,039). Counts <11 are masked; percentages are of the stated denominator; means are shown with SD.

Characteristic	Confirmed BPC-157 patients (N = 1,039)
Notes confirming use per patient, mean (SD)	2.4 (3.2)
Documented use duration, mean (SD), days	74 (198)
Documented supply route, n (%):	
Consumer-obtained	205 (19.7)
Clinician prescription	74 (7.1)
Not documented	760 (73.1)
Documented symptomatic response — of 354 assessed, n (%):	
Improved	279 (78.8)
No change	58 (16.4)
Worse	17 (4.8)
Injury and clinical characterization — among 426 with note-ascertained injury covariates:	
Healthcare utilization, mean (SD), contact-days/yr	15.8 (18.2)
Prior treatments for the injury, mean (SD)	1.61 (1.38)
Current or former smoking, n (%)	50 (11.7)
Injury type, n (%):	
Back disc	123 (28.9)

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

Osteoarthritis	85 (20.0)
Tendinopathy	35 (8.2)
Myalgia	34 (8.0)
Fracture	28 (6.6)
Joint pain	24 (5.6)
Rotator cuff	20 (4.7)
Strain	19 (4.5)
Other MSK	18 (4.2)
Meniscus	14 (3.3)
Ligament	11 (2.6)
Other (each <11)	15 (3.5)
Injury severity, n (%):	
Not stated	251 (58.9)
Severe	76 (17.8)
Moderate	57 (13.4)
Mild	42 (9.9)
Baseline healing status, n (%):	
Chronic non-healing	161 (37.8)
Not stated	121 (28.4)
Acute	75 (17.6)
Post-surgical	46 (10.8)
Subacute	23 (5.4)
Time from injury to treatment, n (%):	
Not stated	216 (50.7)
>12 months	74 (17.4)
<1 month	57 (13.4)
3–12 months	47 (11.0)
1–3 months	32 (7.5)

Table 3. Physician validation of the large-language-model (LLM) ascertainment. Single-site physician review across the network; accuracy is shown with its 95% confidence interval and the number of adjudicated items (n). Every large-language-model adjudication that appears in a figure is referenced in the final column; agreement below 80% is read descriptively, and the post-surgical-region estimate reflects a small reviewed sample (n = 10).

What is the physician expert validating?	Physician-validated accuracy (95% CI)	Corresponding Figure(s)
BPC-157 exposure ascertainment (use vs non-use)	87.2% (74.3–95.2; n = 47)	Figure 1A, 1B
BPC-157 duration-of-use ascertainment	96.7% (88.6–99.1; n = 60)	Figure 2E
BPC-157 first-use-date ascertainment relative to first clinical-note documentation	94.0% (83.8–97.9; n = 60)	Figure 2F
Mono/combination-therapy and co-therapy ascertainment	96.0% (86.3–99.5; n = 50)	Figure 1C, 1D, 4E, 5C–5E
Lifestyle-interventions ascertainment	74.4% (68.1–81.6; n = 50)	Figure 1E
Route-of-administration ascertainment	92.8% (89.4–96.7; n = 50)	Figure 1F
Direct-to-consumer (supply) channel ascertainment	93.8% (82.8–98.7; n = 48)	Figure 1G
Indications-for-use ascertainment	83.9% (67.4–92.9; n = 31)	Figure 3A

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

Injury-site ascertainment	74.1% (61.6–83.7; n = 58)	Figure 3B
Post-surgical-region ascertainment	100% (72.2–100; n = 10)	Figure 3C
Musculoskeletal indication-subtype ascertainment	73.5% (59.7–83.8; n = 49)	Figure 3D
Reasons-for-discontinuation ascertainment	80.6% (67.1–90.2; n = 48)	Figure 3E
Drug-attributed adverse-event ascertainment	93.9% (83.1–98.7; n = 49)	Figure 3F
Any-adverse-event ascertainment	92.0% (80.8–97.8; n = 50)	Figure 3F
Documented symptomatic-response ascertainment (direction; longitudinal trajectory; use-duration gradient)	88.2% (77.1–95.4; n = 43)	Figure 4A, 4D, 5B
Response-reporter ascertainment (patient-reported vs clinician-observed)	93.3% (84.1–97.4; n = 60)	Figure 4B
Type of symptomatic/clinical improvement (benefit-type) ascertainment	78.4% (66.3–87.5; n = 44)	Figure 4C
Quality-of-life and functional-domain ascertainment	83.3% (70.6–92.5; n = 48)	Figure 5A
Outcome-statement-type ascertainment (patient-perceived effectiveness vs product/injection-safety)	85.7% (72.2–93.3; n = 60)	Figure 3G

Using an LLM, reasons for using BPC-157 were ascertained from clinical notes, with a physician-adjudicated extraction accuracy of 84% on manual review (Table 3). A specific documented reason for use was determined for 644 of 1,039 (62%) patients, with some patients having more than one documented reason. The most common indication was pain (338 patients, 33% of the cohort), followed by gastrointestinal conditions including reflux, IBD or gastritis (148, 14%), prior injury (118, 11%), post-surgical healing (107, 10%), and performance enhancement (42, 4%) (Figure 3A). The most common anatomic sites of injury driving BPC-157 use were the shoulder, knee, spine/back, elbow, and hip/pelvis (Figure 3B), with leading conditions including rotator cuff pathology, other tendinous conditions (e.g., tendinitis, tendinopathy, tears), and ligament injuries (Figure 3D); among patients treated after surgery, the shoulder, abdomen or gastrointestinal tract, and knee were the most common operative regions (Figure 3C). On the other hand, among 129 patients with confirmed discontinuation of BPC-157 during the study period, the most commonly cited reasons for discontinuation included advice from a clinician (59 of 129, 46%), adverse effects (24 of 129, 19%), and perceived ineffectiveness (21 of 129, 16%), while issues related to treatment cost or availability were rarely noted (Figure 3E). More granular assessment of the reported adverse effects highlighted neuropsychiatric, dermatologic (including injection site reactions), and gastrointestinal symptoms as the most prevalent (Figure 3F). Physician-adjudicated accuracy for drug-attributed adverse-event ascertainment was 93.9% (Table 3). Documented outcome statements were predominantly descriptions of patient-perceived effectiveness rather than product- or injection-safety observations (physician-adjudicated accuracy 85.7%; Figure 3G, Table 3).

Given the inherent limitations relating to the extent and accuracy of capture of BPC-157 in EHRs, we did not attempt to quantify its effectiveness or compare it to other standard-of-care treatments. Instead, we utilized LLMs to capture qualitative patient-reported and/or clinician-documented outcomes relating to BPC-157 use. Of the 1,039 patient analytic cohort, 354 patients had some form of documented directional symptomatic response, with 279 of 354 (79%) reporting improved symptoms, 16% reporting no change, and 5% reporting worse symptoms (Figure 4A). On manual review, the LLM accuracy in determining the directionality of symptom change was 88% (Table 3). Of note, the

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

vast majority of these symptomatic improvements (263 of 279, 94%) were reported by patients, while 16 (6%) were only observed by the clinician (e.g., on physical examination) (Figure 4B); response-reporter ascertainment had a physician-adjudicated accuracy of 93.3% (Table 3). Among patients with a documented benefit type, pain relief was most common, followed by wound or tissue healing and functional return (Figure 4C). The documented-improvement rate did not differ significantly across strata of documented use-duration (Cochran–Armitage trend P = 0.76; Figure 4D). Among improved patients, roughly one-quarter were on BPC-157 monotherapy while most had co-therapies, most commonly TB-500, testosterone, and NSAIDs (Figure 4E). In an additional analysis, quality-of-life and functional domains — pain interference, mobility, and physical activity — showed documented improvement in the majority of assessable patients (Figure 5A); longitudinal note sentiment was most often consistently improved across notes (Figure 5B); and co-therapy patterns among domain-improved patients mirrored those of the overall improved cohort (Figure 5C–5E).

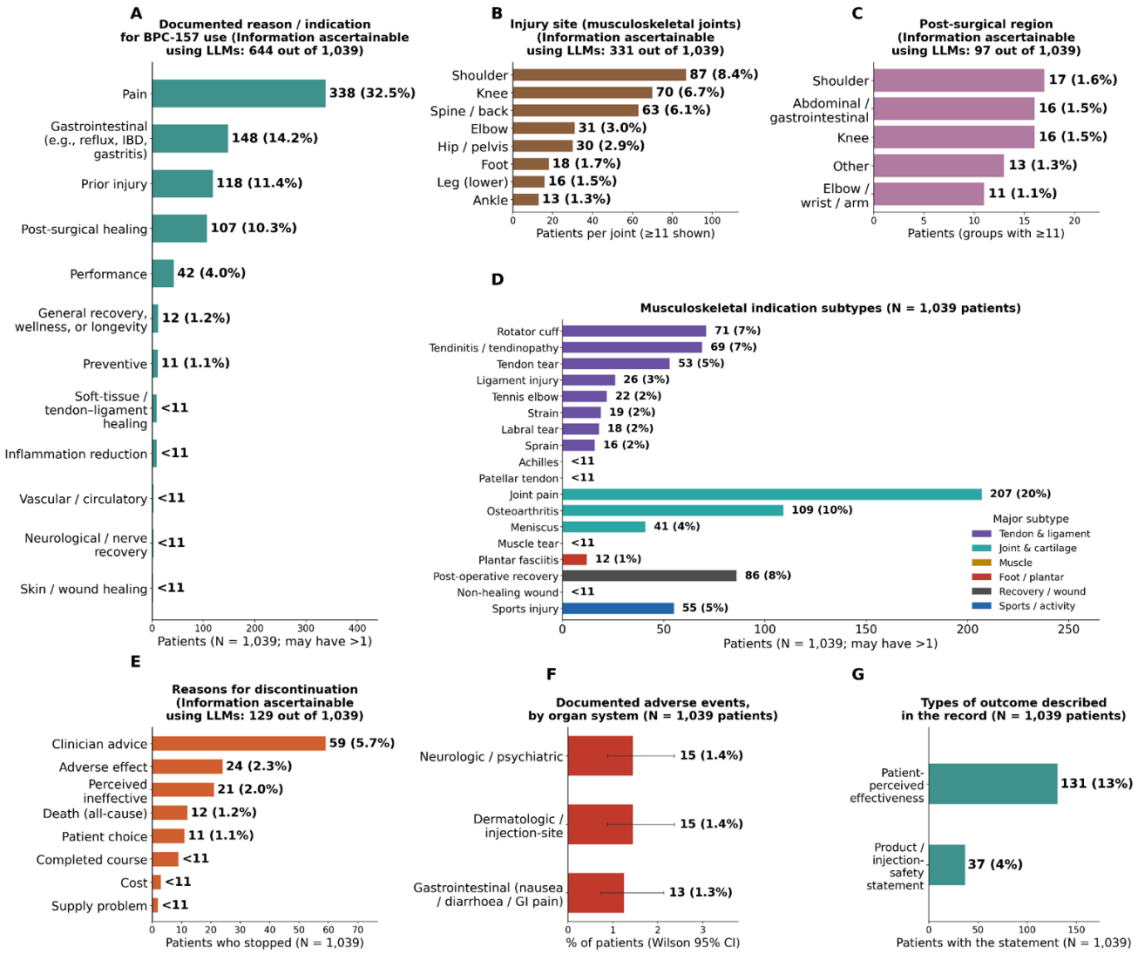


Figure 3. LLM-ascertained documented indications, discontinuation, adverse events, and outcome statements among confirmed BPC-157 patients. Panels A–D, F and G are LLM-ascertained from note text; bars are labelled with the patient count and its percentage of the panel denominator. (A) Documented specific reason/indication categories for use, among patients in whom a specific reason was documented (a patient may have more than one); patients whose BPC-157 use was documented without any specific stated reason are not shown here. (B) Documented injury site by musculoskeletal joint (each patient counted once per joint; joints with ≥11 patients). (C) Post-surgical procedures grouped by anatomical region (groups with ≥11 patients). (D) Musculoskeletal indication subtypes, coloured by major tissue category (see legend). (E) Documented reasons for discontinuation, among patients with a documented stop; the all-cause death category is derived from structured death records rather than note text. (F) Documented adverse events grouped by organ system — neurologic/psychiatric (e.g., mood change, anxiety, insomnia, headache), dermatologic/injection-site (e.g., injection-site reactions, rash), and gastrointestinal (nausea, vomiting) — as a percentage of confirmed patients

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

with Wilson 95% confidence intervals; individual adverse-event terms were each too infrequent (below 11) to display separately. (G) Types of outcome statement documented in the record — whether the note discusses patient-perceived effectiveness or product-/injection-safety — independent of the response direction shown in Figure 4A. The two categories are not mutually exclusive: a patient with both statement types is counted in both bars. Adverse-event findings are documented signals, not prospectively ascertained incidence. Counts below 11 are suppressed.

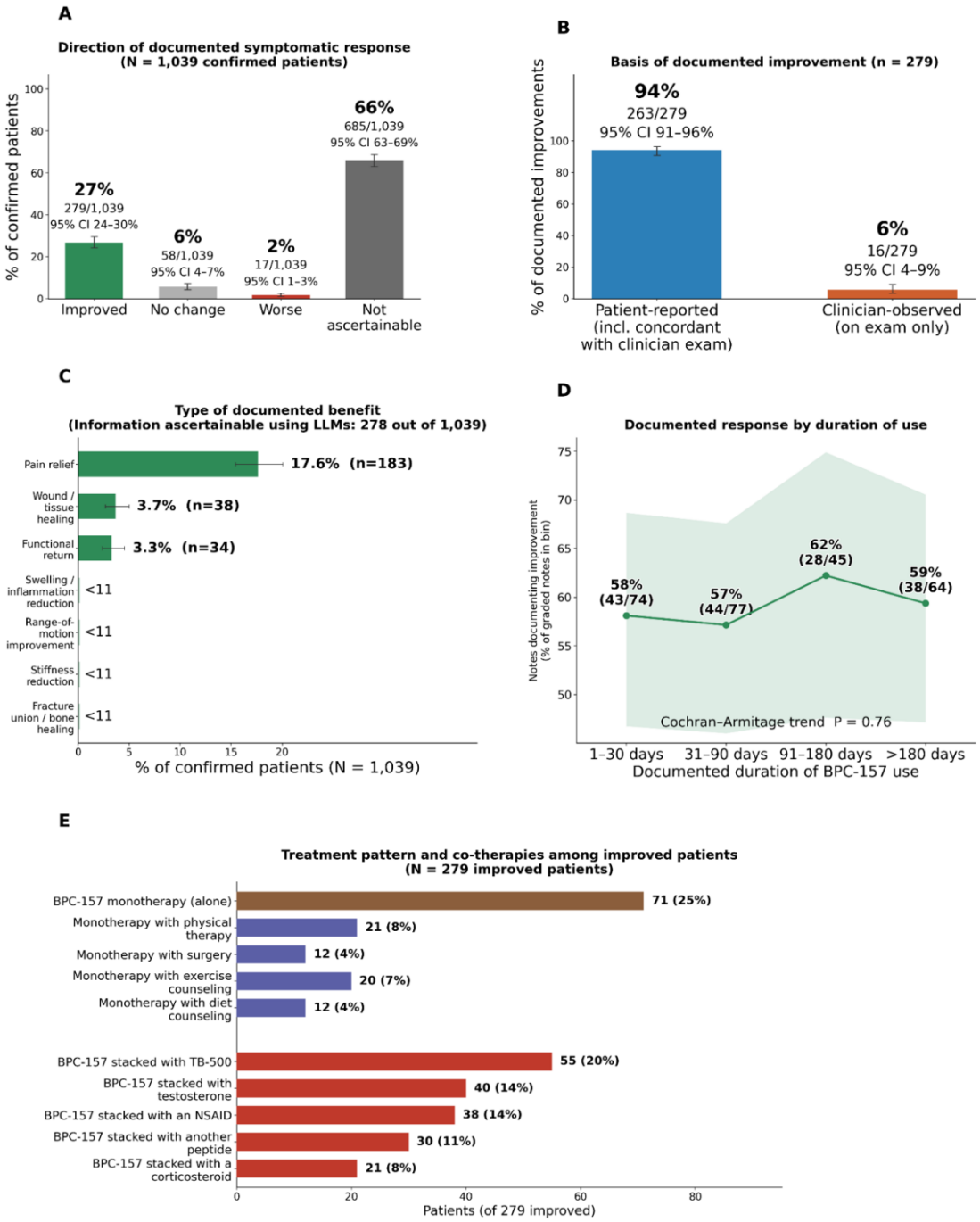


Figure 4. LLM-ascertained documented symptomatic and functional response among confirmed BPC-157 patients. All panels are LLM-ascertained. (A) Direction of the documented symptomatic response as a share of all confirmed patients — improved, no change, worse, or not ascertainable (no response direction documented) — with Wilson 95% confidence intervals; all four categories are shown so the cohort is fully accounted for. (B) Basis of the documented improvement: patient-reported (including reports concordant with the clinician examination) versus clinician-observed on examination; this reflects to whom the note attributes the observation, not a judgment of validity. (C) Type of documented benefit among patients with a stated benefit type. (D)

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

Documented-improvement rate across ordered strata of documented use-duration, tested by the Cochran–Armitage trend test. (E) Treatment pattern among improved patients — BPC-157 monotherapy alone and monotherapy combined with a lifestyle (non-drug) intervention, and, separately, the specific agents BPC-157 was stacked with; bars are a percentage of the improved group, and the stacked-agent bars overlap (a patient may use several). Counts below 11 are suppressed.

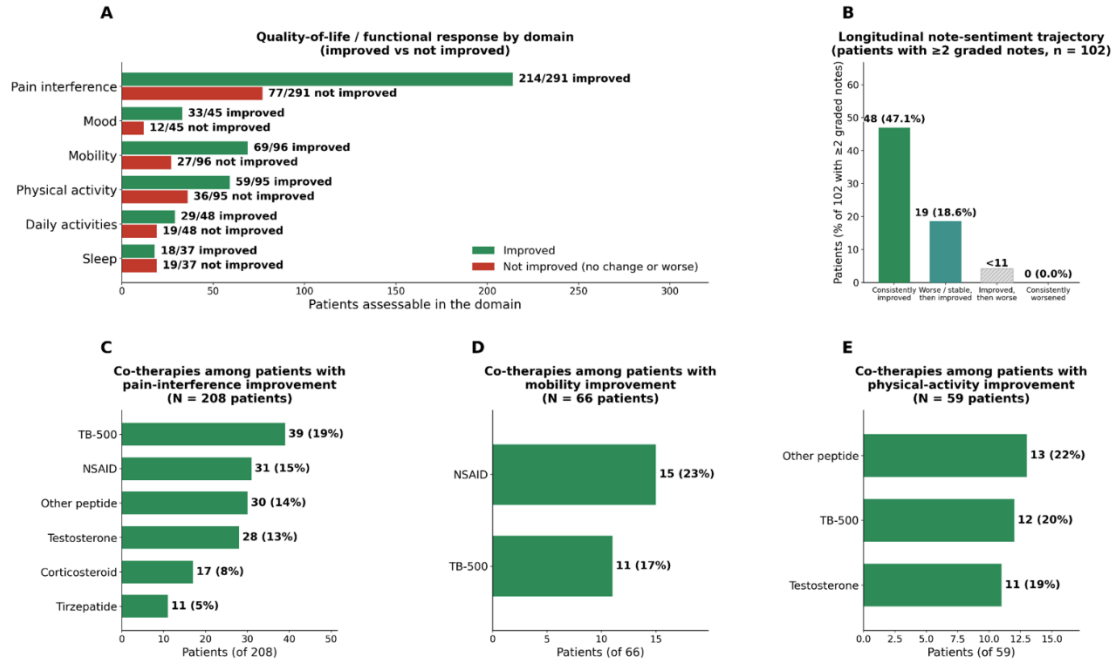


Figure 5. LLM-ascertained quality-of-life and functional response, longitudinal note sentiment, and co-therapy patterns among confirmed BPC-157 patients. All panels are LLM-ascertained from clinical-note text. (A) Quality-of-life and functional response by domain, shown as the number of patients improved versus not improved (no change or worse) out of those assessable in each domain. (B) Longitudinal note-sentiment trajectory among patients with two or more graded notes, classified as consistently improved, initially worse or stable then improved, improved then worsened, or consistently worsened. (C–E) Co-therapies used alongside BPC-157 among patients reporting improvement in pain interference (C), mobility (D), and physical activity (E) — the domains in which at least one co-therapy reached 11 patients — as a percentage of the improved group; co-therapy categories overlap (a patient may use several). Counts below 11 are suppressed.

4. Discussion

There has been a significant increase in public attention around peptides as possible therapeutics in the wake of the recent recommendation by an FDA committee to reclassify multiple peptides (including BPC-157) as Category 1 compounds, which would allow them to be compounded with a valid prescription [24]. While BPC-157 and other peptides have been studied in animal models, assessment in humans remains extremely limited [1,6–8]. Here, we begin to address this gap by performing the first large-scale assessment of BPC-157 utilization and patient-reported outcomes across a large federated network of U.S. health centers. The results highlight a growing number and percentage of patients who report using BPC-157 over the past five years and provide important context around utilization characteristics including sources of access, administration routes, and reasons for initiation and discontinuation. To summarize, in the analyzed cohort of patients reporting use to a healthcare provider, BPC-157 was most commonly initiated to help manage pain or gastrointestinal symptoms, obtained from a compounding pharmacy or gray-market peptide vendor, administered subcutaneously, and discontinued in the setting of clinician advice or adverse effects.

This study was not designed to quantify the effectiveness of BPC-157 in this cohort of real-world users for several reasons, including but not limited to the absence of critical pharmacologic

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

information (e.g., dose and frequency of administration), the likelihood of significant selection bias, and the lack of standardized outcomes documented in the routine care setting. However, we did apply LLMs to extract qualitative reports of changes in symptoms that were attributed to or temporally associated with BPC-157 use. Of the 1,039 patients in the analytic cohort, such reports were present for 354 (34%), with 79% of these patients reporting improvement in symptoms related to BPC-157, 5% reporting worsening, and 16% reporting no change, and manual review confirmed high accuracy of these symptomatic-response extractions (88.2%; Table 3). These findings do not establish that BPC-157 is clinically effective and may be impacted by various confounding factors including multiple sources of bias (e.g., selection, confirmation, and reporting) and placebo effect. However, the confirmation that a reasonable fraction of reporting BPC-157 users have attributed at least some symptomatic improvement to this peptide warrants attention and could motivate prospective randomized studies designed to assess the efficacy of BPC-157 across a range of orthopedic and gastrointestinal conditions. Of course, robust early phase studies to characterize the pharmacodynamic, pharmacokinetic, and tolerability profiles of BPC-157 in humans will be critical to enable such efficacy studies.

The documented reasons for discontinuation of BPC-157 are also enlightening and motivate further directed research. Among the 129 patients with documented discontinuation, the most commonly cited reason was clinician advice. This likely reflects a general discomfort of healthcare providers with their patients using compounds that are not approved for use by the FDA on the basis of randomized controlled trials. The second most commonly cited reason for discontinuation was adverse effects, including neuropsychiatric, dermatologic, and gastrointestinal symptoms. Further analyses of the true incidence and severity of these complications will provide important context in the consideration of whether this peptide should be considered safe for use in humans.

The methods developed and utilized in this study can be useful in an era of increasing use of unapproved compounds. These compounds are inherently difficult to study in the real-world setting due to the lack of structured, standardized information that is typically present for approved and prescribed therapies (e.g., prescriptions containing dose, frequency, and routes of administration and prescription fill data which can serve as a proxy for treatment adherence). Currently, unstructured clinical notes are an essential source of information within the EHR to enable the study of BPC-157 or other similar compounds. The large-scale application of LLMs across deidentified patient records enables rapid curation of relevant clinical data, with manual physician adjudication serving as a critical step to validate the model outputs [11]. As the population of patients who report using BPC-157 or other unapproved compounds continues to increase, the methods employed here can easily scale to allow near real-time assessments of utilization and outcomes [25].

This study has several limitations. First, it is a retrospective study analyzing patient-reported use of and outcomes associated with an unapproved compound. As has been mentioned previously, this framework is susceptible to multiple sources of bias and confounding. In this context, it is important to recognize that any results presented here must be interpreted as descriptive characteristics of a subset of patients who have chosen to reveal their use of BPC-157 to a provider in the captured health care systems. There are certainly additional patients who have used BPC-157 but not reported it, and the group analyzed in this study is likely not completely representative of the overall population of BPC-157 users. Second, the study did not include objective assessments or measures of change in patient symptoms, underlying diseases, or physiology. Future studies incorporating additional longitudinal imaging-based and functional assessments would be useful. Third, there was no active comparator (control) group. Although this was intentional for the reasons outlined previously, it does limit the interpretability of the findings. Fourth, the analytic cohort is relatively small, and the subsets of patients who contribute each data type of interest (e.g., a documented reason for initiation, duration of use) are even smaller. Fifth, the analyses that compare pre- versus post-BPC-157 intervals rely on an imperfect index date, which is set by the first clinical note in which use was confirmed. This date does not align perfectly with (and indeed may be quite different than) the actual date of treatment initiation. Sixth, validation of the various LLM prompts

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

was not performed in a blinded fashion, which could lead to overestimation of model accuracy due to confirmation or anchoring bias. Finally, the study population is drawn from a federated network that is not fully representative of the U.S. population.

With these caveats in mind, this study represents an important first step toward characterizing the patients who have reported use of BPC-157. This patient-centered real-world study documented a considerable fraction of patients using BPC-157 who subjectively experienced improvements in symptoms while taking this compound. The existence of such patient-reported positive outcomes could motivate prospective studies to better define the safety and efficacy of BPC-157 in humans.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Author Contributions: Venky Soundararajan conceived the study. Venky Soundararajan and AJ Venkatakrishnan designed the study and supervised the work. Adhikaar Marwaha performed the analyses with input from Karthik Murugadoss. Santhosh Shiv and Gowtham Varma performed the manual curation and validation. AJ Venkatakrishnan, Venky Soundararajan, Christopher J. Gregg and Adhikaar Marawaha wrote the manuscript with inputs from Matthew Hurchik. Karthik Murugadoss contributed to the analytical and software methodology. All authors reviewed the manuscript and approved the final version.

Funding: No external funding was received for this study.

Acknowledgments: The authors acknowledge the use of the nference federated AI platform. The authors thank Patrick Lenehan for careful review and helpful feedback.

Conflicts of Interest: The authors are employees of nference, inc., which conducts research collaborations with various biopharmaceutical companies whose therapeutic products are included in this study. None of these companies, nor any other nference collaborator, funded, supported, or had any role in the independent study design, data acquisition, analysis, interpretation, manuscript preparation, or the decision to submit this work for publication. All analyses were conducted by the authors using de-identified electronic health record data. The authors declare no additional competing interests.

References

1. Vasireddi N, Hahamyan H, Salata MJ, Karns M, Calcei JG, Voos JE, Apostolakos JM. Emerging Use of BPC-157 in Orthopaedic Sports Medicine: A Systematic Review. HSS J. 2025;21(4):485-495. PMID: 40756949.
2. Chang CH, Tsai WC, Lin MS, Hsu YH, Pang JHS. The promoting effect of pentadecapeptide BPC 157 on tendon healing involves tendon outgrowth, cell survival, and cell migration. J Appl Physiol (1985). 2011;110(3):774-780. PMID: 21030672.
3. Cerovecki T, Bojanic I, Brcic L, Radic B, Vukoja I, Seiwerth S, Sikiric P. Pentadecapeptide BPC 157 (PL 14736) improves ligament healing in the rat. J Orthop Res. 2010;28(9):1155-1161. PMID: 20225319.
4. Hsieh MJ, Liu HT, Wang CN, Huang HY, Lin Y, Ko YS, Wang JS, Chang VHS, Pang JHS. Therapeutic potential of pro-angiogenic BPC157 is associated with VEGFR2 activation and up-regulation. J Mol Med (Berl). 2017;95(3):323-333. PMID: 27847966.
5. U.S. Food and Drug Administration. Evaluation of BPC-157-Related Bulk Drug Substances (BPC-157 [Free Base] and BPC-157 Acetate) for Inclusion on the 503A Bulk Drug Substances List. FDA Briefing Document, Pharmacy Compounding Advisory Committee Meeting, July 23-24, 2026.
6. Lee E, Padgett B. Intra-Articular Injection of BPC 157 for Multiple Types of Knee Pain. Altern Ther Health Med. 2021;27(4):8-13. PMID: 34324435.
7. Lee E, Walker C, Ayadi B. Effect of BPC-157 on Symptoms in Patients with Interstitial Cystitis: A Pilot Study. Altern Ther Health Med. 2024;30(10):12-17. PMID: 39325560.
8. Lee E, Burgess K. Safety of Intravenous Infusion of BPC157 in Humans: A Pilot Study. Altern Ther Health Med. 2025;31(5):20-24. PMID: 40131143.
9. World Anti-Doping Agency. World Anti-Doping Code International Standard: Prohibited List 2026. World Anti-Doping Agency; 2026.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 September 2026 doi:10.20944/preprints202609.0501.v1

10. Harkema H, Dowling JN, Thornblade T, Chapman WW. ConText: An algorithm for determining negation, experiencer, and temporal status from clinical reports. J Biomed Inform. 2009;42(5):839-851. PMID: 19435614.
11. Varma G, Murugadoss K, Kurian ME, Varghese J, Venkatakrishnan AJ, Soundararajan V. Auditable large language model curation of clinical notes refines glucagon-like peptide-1 initiation, persistence ascertainment, and compounding use beyond prescription records. Biol Methods Protoc. 2026;11(1):bpag035.
12. Alsentzer E, Rasmussen MJ, Fontoura R, Cull AL, Beaulieu-Jones B, Gray KJ, Bates DW, Kovacheva VP. Zero-shot interpretable phenotyping of postpartum hemorrhage using large language models. NPJ Digit Med. 2023;6:212.
13. Munzir SI, Hier DB, Oommen C, Carrithers MD. A Large Language Model Outperforms Other Computational Approaches to the High-Throughput Phenotyping of Physician Notes. AMIA Annu Symp Proc. 2024.
14. Murugadoss K, Venkatakrishnan AJ, Soundararajan V. Comparison of Cardiometabolic Benefits and Tolerability in Patients on Sustained Low Doses of Semaglutide or Tirzepatide. Biol Methods Protoc. 2026;bpag048. https://doi.org/10.1093/biomethods/bpag048
15. Venkatakrishnan AJ, Matson R, Murugadoss K, Aman A, Anand D, Soundararajan V. Real-World EHR Signals from a Cohort of Blinded Incretin Trial Participants Motivate Novel Indication Opportunities. Preprints. 2026;202609.0311.v1. https://doi.org/10.20944/preprints202609.0311.v1
16. Wilson EB. Probable inference, the law of succession, and statistical inference. J Am Stat Assoc. 1927;22(158):209-212.
17. Cochran WG. Some methods for strengthening the common χ2 tests. Biometrics. 1954;10(4):417-451.
18. Armitage P. Tests for linear trends in proportions and frequencies. Biometrics. 1955;11(3):375-386.
19. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B. 1995;57(1):289-300.
20. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-1457. PMID: 18064739.
21. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015;12(10):e1001885. PMID: 26440803.
22. Murugadoss K, Rajasekharan A, Malin B, Agarwal V, Bade S, Anderson JR, Ross JL, Faubion WA Jr, Halamka JD, Soundararajan V, Ardhanari S. Building a best-in-class automated de-identification tool for electronic health records through ensemble learning. Patterns. 2021;2(6):100255.
23. OpenAI. GPT-4o System Card. Published August 8, 2024. arXiv:2410.21276.
24. Eglovitch JS. FDA advisory committee backs two controversial peptides. Regulatory Focus (Regulatory Affairs Professionals Society). Published July 23, 2026. https://www.raps.org/resource/fda-advisory-committee-backs-two-controversial-peptides.html
25. Murugadoss K, Venkatakrishnan AJ, Soundararajan V. Accelerating Use of Unapproved Retatrutide Is Associated with Weaker Weight Loss and Increased Cardiovascular Symptoms. Preprints. 2026;202608.1193.v1. https://doi.org/10.20944/preprints202608.1193.v1

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.