

Transforming Multi-Clinic Naturopathic EHR Data into a Research Database: A Retrospective Feasibility Pilot Study

Paul Kalnins

pkalnins@nunm.edu

National University of Natural Medicine

Ryan Bradley

University of California San Diego

Steve Chamberlin

Oregon Health & Science University

Michelle Simon

Institute for Natural Medicine

Erin Sweet

Bastyr University

Carlo Calabrese

National University of Natural Medicine

Research Article

Keywords: naturopathic medicine, electronic health records, FHIR, de-identification, natural language processing, health informatics, multi-site data integration, real-world data, practice-based research network

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10348913/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Electronic health records (EHRs) from community-based naturopathic clinics represent a valuable resource for practice-based research. However, no standardized infrastructure exists for aggregating and standardizing these data across clinical sites. Further, analyzing unstructured data from EHRs remains a challenge. This methods-feasibility pilot developed a Health Insurance Portability and Accountability Act (HIPAA)-compliant, cloud-based naturopathic EHR research database with infrastructure supporting multi-site data aggregation, de-identification, and free-text treatment extraction. The pilot evaluated the database's utility through patient population characterization, treatment pattern analysis, and proof-of-concept analyses of clinician-ordered treatments and longitudinal trajectories within hemoglobin A1c (HbA1c)-defined cohorts.

Methods

Retrospective EHR data were extracted via the Fast Healthcare Interoperability Resources (FHIR) API from seven community naturopathic clinics in Washington State and were loaded into a structured cloud-based research database. Records were de-identified using automated protected health information detection and confirmed by a manual audit. Clinician-ordered treatments were extracted from clinical notes, categorized into 12 treatment modalities and three practice metrics, and validated through manual review of a stratified random sample (positive predictive value [PPV] 93.3%, 95% CI: 90–96%). HbA1c-based subgroups were analyzed as exemplar use cases for evaluating treatment extraction and longitudinal signal detection, using paired t-tests (first vs. last HbA1c) within glycemic subgroups.

Results

The assembled database included 58,498 patients (mean age 42.7 years; 68.1% female), 469,162 clinical encounters, 858,394 clinical notes, and 3.6 million observations (labs/vitals) spanning 18 years (2008–2026). Common documented conditions included fatigue (29.8%), vitamin D deficiency (9.6%), and insomnia (8.0%). In a cohort with elevated HbA1c ($n = 2,864$), the most frequently documented treatment modalities were nutritional supplements (80.1%), diet/nutrition counseling (47.3%), and botanical medicine (43.6%). Longitudinal HbA1c analyses ($n = 5,576$) showed significant declines in both diabetes and prediabetes subgroups (both $p < 0.001$), demonstrating the database's capacity to detect clinically meaningful change over time.

Conclusions

This pilot demonstrates the feasibility of transforming multi-clinic naturopathic EHR data into a queryable research database using a HIPAA-compliant cloud-based pipeline with validated treatment extraction capabilities. The resulting infrastructure supports diverse analytic applications, providing a scalable foundation for practice-based research across complementary and integrative health clinical settings.

Introduction

Naturopathic physicians (NDs) provide integrative, patient-centered care that emphasizes individualized, multimodal management of chronic conditions¹ and have been described as whole health specialists.² Despite growing patient utilization in the United States, systematic, large-scale data on community-based naturopathic practice remain limited, constraining health services and outcomes research in this field.^{3,4} Advancing the evidence base requires research infrastructure capable of systematically characterizing real-world clinical practice and supporting longitudinal outcome evaluation.

Prior research has taken three main approaches, each with important limitations. Practice characterization studies have relied primarily on practitioner surveys and administrative billing data. Surveys capture what clinicians report prescribing but not what is documented in patient records.⁵ Administrative datasets—while enabling large-scale analyses of diagnoses and visit patterns⁶—do not capture individualized treatment plans that are usually recorded in unstructured clinical notes, which is the primary location where naturopathic therapeutic decision-making is documented. Outcomes research has been similarly constrained to small chart reviews, protocol-based studies in academic settings, and systematic reviews of this same small, heterogeneous evidence base, with limited generalizability to community practice.^{7,8,9} The Naturopathic Patient Database (NPD) demonstrated that a structured, queryable database could capture individualized treatment protocols in academic teaching clinics,¹⁰ but its reliance on manual data entry and restriction to academic settings limits scalability and applicability to the independent community practices where most naturopathic care is delivered. Extending this model to community practice would be greatly facilitated by automated approaches capable of extracting and standardizing data directly from routine clinical documentation across multiple clinics.

Practice-based research networks (PBRNs) have addressed analogous gaps in primary care and other disciplines by pooling real-world electronic health records (EHR) data across independent practices into shared infrastructure for multi-site research.¹¹ No equivalent multi-site EHR-based research network exists for naturopathic medicine. Recent adoption of cloud-based EHR platforms by community ND clinics, combined with the emergence of the HL7 Fast Healthcare Interoperability Resources (FHIR) standard, has made automated multi-clinic data aggregation increasingly feasible.¹² The central informatics challenge is that the most clinically distinctive aspects of naturopathic care are often embedded in unstructured clinical narratives rather than discrete coded fields, requiring infrastructure capable of extracting, de-identifying, and structuring this content at scale.¹³

This methods-feasibility pilot had four objectives: (1) to develop an automated, FHIR-based data extraction pipeline capable of ingesting multi-clinic naturopathic EHR data; (2) to de-identify extracted records and assemble them into a consolidated, Health Insurance Portability and Accountability Act (HIPAA)-compliant cloud-based research database spanning seven community clinics; (3) to characterize the assembled patient population, clinical content, and documented treatment patterns; and (4) to demonstrate the analytic utility of the resulting database through analyses of longitudinal HbA1c trajectories and clinician-ordered treatment patterns within glycemically defined patient cohorts.

Methods

Study design and setting

This retrospective, observational pilot study assessed the feasibility of constructing a HIPAA-compliant, secure cloud-based research database by extracting and de-identifying EHR data from seven community-based naturopathic clinics in Washington State. Five of the clinics participate in the Institute for Natural Medicine (INM) naturopathic residency network. All clinics used CharmHealth, a cloud-based EHR system with a standardized application programming interface (API) supporting the FHIR R4 standard.¹⁴

The study was approved by the National University of Natural Medicine (NUNM) Institutional Review Board (#CK021726). Participating clinics executed data sharing agreements, and researchers signed data use agreements prior to data access. Data were exported from the CharmHealth data center in May 2026, capturing all available records from each clinic's inception through the export date (approximate range: 2008–2026). Because we were fundamentally evaluating feasibility, no *a priori* sample size calculation was performed. All available records from participating clinics were included.

Pipeline overview

The data pipeline operated in four sequential stages, each described in detail in the subsections below. Raw EHR data were first exported from each clinic via the CharmHealth FHIR R4 API and stored in a HIPAA-compliant cloud environment. Data were then de-identified through a multi-stage process combining automated protected health information (PHI) detection, rule-based free-text redaction, and a manual safety audit. De-identified data were subsequently loaded into a structured relational research database (Google BigQuery) and standardized across all seven clinics. Finally, clinical notes were processed through a rule-based clinical text extraction pipeline to extract and classify clinician-ordered treatments into 12 treatment categories and three clinical practice metrics, with results validated by manual review.

Data extraction

Structured and unstructured EHR data were exported for each clinic via the CharmHealth API. Extracted data categories included patient demographics, encounters, diagnoses, observations, clinical notes, care plans, medication orders, and practitioner resources. All data were stored in a Google Cloud Platform (GCP) environment operating under a HIPAA Business Associate Agreement. Access was restricted via Identity and Access Management (IAM) role assignments. Data were encrypted at rest (AES-256) and in transit (TLS), with activity logged throughout the PHI-containing phase via Cloud Audit Logs. Clinic participation was facilitated through a self-service web application that guided clinic administrators through API authorization and credential setup.

De-identification

De-identification was performed in multiple stages:

Stage 1 — Structured field removal and date shifting. All 18 categories of protected health information (PHI) defined under the HIPAA Privacy Rule¹⁵ were addressed using the Google Cloud Healthcare API.¹⁶ Direct identifiers (names, contact information, and addresses) were removed from structured data fields. Rather than removing dates entirely, dates were shifted by a randomly assigned, patient-specific offset (± 0 –364 days) to preserve within-patient time intervals for longitudinal analysis. Because date shifting does not satisfy HIPAA Safe Harbor’s date-removal requirement, the dataset was treated as a Limited Dataset under 45 CFR § 164.514(e), with dates retained as permitted and access restricted to approved investigators under executed data use agreements. Internal record identifiers were replaced with anonymous unique identifiers. Patient birth year was retained, and patients born before 1936 were recorded as “ ≤ 1935 ” per HIPAA requirements.

Stage 2 — Free text redaction. Clinical notes, care plan text, and observation fields were processed using the Google Cloud Healthcare API (powered by Cloud Data Loss Prevention). Detected PHI tokens were replaced with labeled placeholders. Specifically, person names were pseudonymized via crypto-hash replacement, dates were date-shifted using the same patient-specific offset applied in Stage 1, and contact identifiers (email addresses, phone numbers, street addresses, Social Security numbers [SSNs], medical record numbers, driver’s license numbers, URLs, and IP addresses) were replaced with labels.

Stage 3 — Additional rule-based redaction. Multiple iterative passes were applied to free-text fields to address PHI categories not reliably detected by automated scanning. Person names were detected using spaCy (an open-source Python NLP library) named-entity recognition (NER). Regular expression pattern rules were applied to identify provider names, clinic names and addresses, employer names, and social history references that could enable re-identification.

Manual safety audit. Following de-identification, a custom audit tool was applied to all de-identified records, checking for: residual structural identifiers (patient name, contact information, and address fields still present after de-identification); out-of-range postal codes and implausible dates; and PHI patterns in free text via regular expression matching (SSNs, phone numbers, email addresses, date

formats, medical record number (MRN)-like strings, and person names). Flagged items were triaged by severity and manually reviewed, with all being confirmed as non-PHI before data were loaded into the research database. The audit tool is publicly available in the project repository (see Data Availability statement).

PHI data lifecycle. Following successful audit, all original raw data files and intermediate PHI-containing data stores were deleted within 30 days of audit completion.

Database construction

De-identified data were loaded into a structured research database (Google BigQuery). Key processing steps included retaining patient birth year from source records, standardizing laboratory test codes, extracting diagnosis codes, and converting clinical notes from formatted to plain text. The final database, residing in BigQuery, comprises 11 data tables covering all extracted data categories, plus a data dictionary describing each field.

Clinical text structuring and classification validation

Prior to cohort analysis, de-identified clinical notes were parsed into a queryable treatment table. Notes were segmented into SOAP (subjective, objective, assessment, plan) sections using pattern matching on common section header strings (e.g., “Plan:”, “Assessment & Plan:”, “Patient Instructions:”). This approach was applied uniformly across all seven clinics without site-specific adjustments. Although CharmHealth provides a consistent FHIR R4 API schema, note formatting varies by clinic and provider — some using structured templates, others using free-form narrative — and this heterogeneity introduces variability in extraction performance (see Limitations). Extraction targeted plan and patient instruction sections only. Subjective sections were excluded to focus on clinician recommendations rather than patient-reported history; assessment sections were excluded because keyword matching cannot reliably distinguish clinician-ordered interventions from patient-reported behaviors in that context.

A curated keyword dictionary was developed to identify and classify clinician-ordered treatments — the therapeutic recommendations and prescriptions recorded in clinical care plans — into 12 categories reflecting the treatment types most frequently documented across the clinical note corpus: (1) nutritional supplements (vitamins, minerals, essential fatty acids, probiotics, amino acid derivatives), (2) botanical medicine (plant-based preparations, including extracts), (3) diet/nutrition, (4) exercise/physical activity, (5) stress/mindfulness, (6) sleep hygiene, (7) physical therapy/bodywork, (8) TCM/acupuncture, (9) IV therapy, (10) hydrotherapy, (11) homeopathy, and (12) hormone therapy (non-prescription hormones such as DHEA and prescription hormones). Three additional categories (lab and diagnostic ordering, specialist referral, and substance cessation counseling) were extracted using the same approach but reported as clinical practice metrics rather than clinician-ordered treatments.

Keywords were developed iteratively through five rounds of clinician review. Treatment mentions were classified as clinician orders only when phrasing was imperative or conditional (e.g., “take,” “continue,” or “consider”). Patient-reported and historical phrasings were excluded through SOAP-section filtering and phrasing-context rules that distinguished clinician recommendations from patient-reported behaviors. All reported frequencies therefore represent clinician-documented orders rather than patient adherence. The complete keyword dictionary is available in the project repository (see Data Availability statement).

Text classification method selection. Keyword matching was selected as the treatment extraction method based on complete transparency (every rule is explicit and clinician-reviewable) and sufficient precision confirmed by manual validation (PPV 93.3%).

Validation. Treatment category assignments were validated by manual review of 300 randomly sampled treatment rows, stratified at 20 rows per category. Two investigators independently reviewed all rows; 15 non-evaluable rows were excluded from inter-rater analysis, leaving 285 paired ratings. Overall PPV was 93.3% (95% CI: 90–96%, Wilson score interval), ranging from 80% (nutritional supplements, exercise/physical activity) to 100% (botanical medicine, homeopathy, hormone therapy, hydrotherapy, stress/mindfulness). Observed inter-rater agreement was 95.4% (Cohen’s $\kappa = 0.66$; Prevalence and Bias Adjusted Kappa [PABAK] = 0.91); the lower κ reflects a well-documented prevalence effect in datasets with high true-positive rates,¹⁷ and PABAK is reported as the primary agreement statistic.

Statistical and analytic methods

Cohort definitions and participant flow. Two illustrative analyses were conducted to demonstrate the analytic utility of the assembled database: clinician-ordered treatment type and frequency, and longitudinal HbA1c trajectory analysis. HbA1c was selected as an exemplar use case because it was both clinically meaningful and among the most consistently available longitudinal laboratory measures across participating clinics. HbA1c measurements were identified via a two-step approach: records with standard Logical Observation Identifiers Names and Codes (LOINC) identifiers were identified by code matching, and records exported without codes were identified by keyword matching on test name strings (e.g., “hemoglobin A1c,” “HbA1c,” “glycated hemoglobin”). Elevated HbA1c was defined as any reading in the prediabetes ($\geq 5.7\%$ to $< 6.5\%$) or diabetes ($\geq 6.5\%$) range.

Two analytic cohorts were derived from the full database (N = 58,498). The **Elevated HbA1c Treatment Cohort** (n = 2,864) comprised patients with ≥ 1 elevated HbA1c reading ($\geq 5.7\%$) and ≥ 1 documented clinician-ordered treatment. The **Longitudinal HbA1c Cohort** (n = 5,576) comprised all patients with ≥ 2 HbA1c readings on distinct dates, regardless of baseline glycemic category; first and last readings were used for analysis. For stratified analyses, this cohort was divided into three subgroups by baseline glycemic status: normoglycemic ($< 5.7\%$, n = 3,994), prediabetes (5.7% to $< 6.5\%$, n = 1,344), and diabetes ($\geq 6.5\%$, n = 238).

HbA1c longitudinal analysis. For the Longitudinal HbA1c Cohort, longitudinal HbA1c change was assessed by paired t-test comparing first and last available HbA1c readings per patient, stratified by baseline glycemic category: diabetes ($\geq 6.5\%$), prediabetes (5.7% to $< 6.5\%$), and normoglycemic ($< 5.7\%$). First-versus-last analysis was selected because the objective was demonstration of longitudinal signal detection rather than estimation of treatment effects. The proportion of patients with any improvement in HbA1c (last reading $<$ first reading) was also calculated. Three sensitivity analyses were conducted: (1) excluding the largest contributing clinic; (2) restricting to patients with ≥ 90 and ≥ 180 days between first and last readings; and (3) a linear mixed-effects model with clinic as a random intercept and glycemic group as a fixed effect. In this model, first-to-last HbA1c change served as the outcome variable; time is therefore embedded in the outcome measure rather than included as a separate continuous predictor, and the model addresses clinic-level clustering but does not model individual trajectories over time.

Treatment type, frequency, and co-occurrence. Treatment type and frequency were summarized descriptively for the Elevated HbA1c Treatment Cohort and across all patients with any HbA1c reading, stratified by baseline glycemic category. Frequencies were expressed as the percentage of patients with at least one documented order in each of the 12 treatment categories. Pairwise co-occurrence was quantified using the Jaccard similarity index, calculated as the number of patients with documented orders for both treatments divided by the number with orders for at least one of the two; values range from 0 (no co-occurrence) to 1 (complete overlap).

Cohort construction and descriptive analyses were performed in the research database (BigQuery). Paired t-tests and mixed-effects models were performed in Python 3.13 (scipy 1.17, pandas 3.0, statsmodels 0.14). Figures were generated using matplotlib 3.10. Statistical significance was set at $p < 0.05$ (two-tailed). No correction for multiple comparisons was applied, consistent with the descriptive and exploratory nature of these analyses. Missing data were not imputed and each analysis used complete cases for the variable(s) of interest.

Results

Database characteristics

The assembled database contains 58,498 patients across seven clinics (five Institute for Natural Medicine residency network clinics, two non-network ND practices), with 469,162 encounters, 858,394 clinical notes, and 3.6 million structured observations (laboratory results and vital signs) spanning 2008–2026. Clinical note extraction yielded 642,858 treatment-related records across 12 treatment categories and three clinical practice metrics (lab/diagnostic ordering, specialist referrals, and substance cessation counseling). Table 1 summarizes patient demographics and health characteristics.

Table 1
Patient demographics and health characteristics

Characteristic	%, Mean, or Median (IQR), n
Demographics — all patients (N = 58 498)	
<i>Sex (birth sex)</i>	
Female	68.1%
Male	30.6%
Unknown	1.3%
<i>Age^a</i>	
Mean age, all (yr)	42.7
Mean age, Female (yr)	43.6
Mean age, Male (yr)	40.3
Age 0–17	11.9%
Age 18–34	18.7%
Age 35–49	32.7%
Age 50–64	22.8%
Age 65+	12.8%
<i>Race^d</i>	
White	9.3%
Other (reported)	1.7%
Unknown / not reported	89.1%
Health characteristics — adult patients (age ≥ 18)^b	
<i>Anthropometrics</i>	
Height (cm)	167.6 (162.6–175.3), n = 16 039
^aAges derived from birth year and are approximate following de-identification date shifting. Age percentages exclude patients with missing birth year (range 0.1%–4.6%). ^bValues are medians (IQR) for the first recorded measurement per patient; n indicates adult patients with at least one recorded value. Coverage varied considerably (vital signs 30–48% of adult patients; laboratory values 13–25%); see Results for explanation. Plausibility filters applied (height 122–244 cm, systolic BP 60–300 mmHg, diastolic BP 30–180 mmHg). ^cBMI computed from height and weight; stored BMI values were sparse and inconsistent across clinics. ^dRace data largely unavailable (89.1% unknown/not reported), reflecting provider non-documentation rather than patient non-disclosure	

Characteristic	%, Mean, or Median (IQR), n
Weight (kg)	72.6 (61.2–86.6), n = 16 890
BMI (kg/m ²) ^c	25.2 (22.1–29.2), n = 15 237
<i>Vital signs</i>	
Systolic BP (mmHg)	115.0 (106.0–124.0), n = 24 388
Diastolic BP (mmHg)	72.0 (64.0–80.0), n = 24 382
<i>Glycemic markers</i>	
Glucose (mg/dL)	90.0 (84.0–96.0), n = 6415
HbA1c (%)	5.4 (5.2–5.6), n = 12 538
<i>Lipid panel</i>	
Total Cholesterol (mg/dL)	194.0 (169.0–222.0), n = 11 633
Triglycerides (mg/dL)	85.0 (62.0–124.0), n = 11 630
LDL-C (mg/dL)	115.0 (93.0–140.0), n = 11 602
HDL-C (mg/dL)	58.0 (48.0–71.0), n = 11 618
<i>Inflammation</i>	
hs-CRP (mg/L)	1.03 (0.50–2.51), n = 9122
<i>Thyroid</i>	
TSH (mIU/L)	1.64 (1.15–2.29), n = 15 477
^a Ages derived from birth year and are approximate following de-identification date shifting. Age percentages exclude patients with missing birth year (range 0.1%–4.6%). ^b Values are medians (IQR) for the first recorded measurement per patient; n indicates adult patients with at least one recorded value. Coverage varied considerably (vital signs 30–48% of adult patients; laboratory values 13–25%); see Results for explanation. Plausibility filters applied (height 122–244 cm, systolic BP 60–300 mmHg, diastolic BP 30–180 mmHg). ^c BMI computed from height and weight; stored BMI values were sparse and inconsistent across clinics. ^d Race data largely unavailable (89.1% unknown/not reported), reflecting provider non-documentation rather than patient non-disclosure	

The database includes 157 practitioner records across seven clinics (range: 3–39 per clinic). Among the 40 (25.5%) with credentials documented, naturopathic doctors were the predominant type (n = 30), including three dual-licensed ND/LACs, with small numbers of allied health professionals also represented (registered dietitian, advanced practice nurse, lactation consultant, physical therapist, and licensed clinical social worker). Practitioner registries in CharmHealth reflect all individuals with EHR system access and likely include inactive or former providers.

The most prevalent conditions among adult patients (age ≥ 18 ; $n = 50,936$) were fatigue (29.8%), vitamin D deficiency (9.6%), insomnia (8.0%), hyperlipidemia (6.3%), and anxiety (6.2%), followed by essential fatty acid deficiency (6.1%), hypothyroidism (5.8%), constipation (5.4%), low libido (3.9%), and prediabetes (3.7%). Patient-level deduplication was applied to all prevalence estimates to minimize inflation from problem-list carry-forward (see Limitations).

Prescription medication data encompassed 46,098 orders. Prescription hormone agents (thyroid and sex hormones) were the most frequently prescribed drug class (4.3% of all patients), followed by antibiotics (2.7%). Low-dose naltrexone was prescribed to 0.6% of patients. Rates of documented antihyperglycemic prescription orders (including metformin, GLP-1 receptor agonists, and insulin) stratified by baseline glycemic severity are presented in Treatment Patterns below.

Documented clinician-ordered treatment frequencies varied meaningfully across clinics. Nutritional supplement orders were the most prevalent category across all sites (range: 39.5%–74.8%), with one site showing an exceptionally high rate of IV therapy (56.9%), another with markedly higher physical therapy/bodywork and stress/mindfulness rates, and a nutrition-focused site with the highest diet/nutrition documentation rate.

HbA1c cohort

Of 58,498 patients, 3,149 (5.4%) had at least one elevated HbA1c reading ($\geq 5.7\%$). Of these, 2,864 (90.9%) had at least one documented clinician-ordered treatment and constituted the Elevated HbA1c Treatment Cohort; the remaining 285, with no clinician-ordered treatment recorded, were excluded. Separately, 5,576 patients with at least two HbA1c readings formed the Longitudinal HbA1c Cohort, of whom 3,994 (71.6%) had a normoglycemic baseline ($< 5.7\%$). Median follow-up ranged from 487 to 642 days across glycemic subgroups (IQR: normoglycemic 379–1,095 days; prediabetes 285–848 days; diabetes 278–879 days).

HbA1c longitudinal findings

The database supported longitudinal stratified analyses of repeated HbA1c measurements across three clinically defined glycemic subgroups (Table 2). Mean HbA1c declined in both elevated groups: diabetes and prediabetes patients each showed reductions from first to last measurement (both $p < 0.001$), while the normoglycemic group was essentially stable (mean change $+ 0.004\%$; $p = 0.28$). These declines are consistent with the database's capacity to detect longitudinal signal, but should not be interpreted as evidence of treatment effectiveness. Regression to the mean, spontaneous glycemic variation, and unobserved concurrent conventional treatment could each contribute to the observed patterns.

Sensitivity analyses were consistent with the primary findings. Excluding the largest contributing clinic yielded directionally identical results in both elevated subgroups (mean HbA1c change: diabetes $- 0.807\%$, prediabetes $- 0.111\%$; both $p < 0.001$). Minimum follow-up restrictions (≥ 90 and ≥ 180 days) produced negligible changes in estimates. A linear mixed-effects model with clinic as a random intercept yielded nearly identical estimates (mean HbA1c change: prediabetes $- 0.122\%$, 95% CI $- 0.157\%$ to $-$

0.087%; diabetes – 0.763%, 95% CI – 0.822% to – 0.705%; both $p < 0.001$), with negligible clinic-level variance in the random effect ($\sigma^2=0.001$).

Figure 1 displays mean HbA1c at six intervals from baseline by glycemic group. In the prediabetes group, mean HbA1c moved from 5.87% to approximately 5.76% by 6–12 months without evidence of upward drift, though attrition bias at longer intervals cannot be excluded. The diabetes group displayed the largest early signal, with greater variability at extended follow-up consistent with the smaller sample size.

Table 2
HbA1c longitudinal outcomes by baseline glycemic category

Baseline Group	n	Mean First HbA1c	Mean Last HbA1c	Mean Change	95% CI	% Improved	Median Follow-up, days [IQR]
Normoglycemic (< 5.7%)	3994	5.28%	5.28%	+ 0.004%	(– 0.004%, + 0.012%)	41.1%	642 [379–1095]
Prediabetes (5.7–<6.5%)	1344	5.87%	5.73%	–0.131%	(– 0.148%, – 0.114%)	67.5%	487 [285–848]
Diabetes (≥ 6.5%)	238	7.97%	7.21%	–0.766%	(– 0.958%, – 0.573%)	69.3%	543 [278–879]

$p < 0.001$ for diabetes and prediabetes groups (two-tailed paired t-test, first vs. last HbA1c); $p = 0.28$ (not significant) for normoglycemic group. 95% CIs computed from the SD of individual paired differences ($1.96 \times SE$). Follow-up is right-skewed; median and interquartile range (IQR) are reported. Total $n = 5576$.

Figure 1. Mean HbA1c trajectory by baseline glycemic group

Mean HbA1c (%) at six time intervals from each patient’s baseline HbA1c reading, stratified by glycemic category (normoglycemic < 5.7%, prediabetes 5.7–<6.5%, diabetes ≥ 6.5%). Shaded bands represent ± 1 SE. Sample sizes at each interval are shown below the x-axis. Dashed horizontal lines indicate the prediabetes (5.7%) and diabetes (6.5%) diagnostic thresholds. Longitudinal HbA1c Cohort: $n = 5,576$ patients with ≥ 2 HbA1c readings on distinct dates.

Treatment patterns

In the Elevated HbA1c Treatment Cohort ($n = 2,864$), nutritional supplement orders were documented for 80.1% of patients, diet/nutrition counseling for 47.3%, botanical medicine for 43.6%, and exercise for 39.6%. Hormone therapy orders (bioidentical/compounded formulations combined with prescription

hormone records) were documented for 21.8% of patients. IV therapy was documented more frequently in patients with normoglycemic baseline HbA1c (24.5%) than in the prediabetes (17.8%) and diabetes (21.1%) groups. Antihyperglycemic prescription drug orders increased markedly with glycemic severity (normoglycemic 1.3%, prediabetes 3.3%, diabetes 12.4%).

Within nutritional supplements, vitamins were most commonly ordered (70.2%, with vitamin D as the dominant agent), followed by minerals (41.9%, dominated by magnesium) and omega-3 fatty acids (34.8%). Within botanical medicine, berberine was the most frequently documented agent (17.3%), classified as botanical medicine given its predominant use in herbal combination products. Ginger (8.4%), garlic (7.2%), adaptogenic formulas (7.0%), milk thistle (6.7%), and curcumin/turmeric (6.3%) were also commonly documented.

Among prescription drugs, prescription hormone agents were most commonly prescribed (8.5%), followed by antidepressants (4.2%), antihypertensives (3.0%), metformin (2.8%), and GLP-1 receptor agonists (1.9%). Insulin was prescribed to 0.5% of patients. Other antihyperglycemic classes (sulfonylureas, SGLT2 inhibitors, DPP-4 inhibitors) were each documented in ≤ 5 patients. Lab and diagnostic ordering was documented for 67.5% of the cohort, specialist referral for 45.3%, and substance cessation counseling for 2.1%.

Treatment co-occurrence. Figure 2 presents a pairwise Jaccard co-occurrence heatmap for the 12 treatment categories plus antihyperglycemic drugs. Nutritional supplements, diet/nutrition, botanical medicine, and exercise formed the most tightly co-ordered cluster. Physical therapy/bodywork and stress/mindfulness showed intermediate co-occurrence with this cluster and with each other. Specialty modalities (IV therapy, acupuncture/TCM, homeopathy, hydrotherapy) showed low co-occurrence with the core cluster and with each other. This three-tier structure — a lifestyle/nutritional core, adjunctive physical and behavioral interventions, and independently used specialty modalities — mirrors the naturopathic therapeutic order, which prioritizes least-force, self-care, and prevention-oriented approaches before drug and surgical interventions.¹⁸

Figure 2. **Pairwise treatment co-occurrence — elevated HbA1c cohort**

Pairwise Jaccard similarity matrix for 12 categories of clinician-ordered treatments plus antihyperglycemic drugs in the Elevated HbA1c Cohort (n = 2,864). Jaccard similarity = (patients with orders for both treatments) / (patients with orders for either treatment); values range from 0 to 1.0. Diagonal entries are 1.0 by definition. Treatment categories are ordered by prevalence. Hormone therapy combines bioidentical/compounded formulations from clinical notes with prescription hormone records.

Discussion

Infrastructure contribution

This pilot demonstrates the feasibility of creating a HIPAA-compliant, multi-clinic naturopathic research database from routine EHR data. The pipeline successfully integrated data from seven independent clinics with diverse practice profiles, patient volumes, and documentation styles, addressing a longstanding data infrastructure gap in naturopathic research. Although all participating clinics used CharmHealth, the pipeline is built on the FHIR R4 standard¹⁴ and is adaptable to other compatible EHR systems. Unlike prior naturopathic database initiatives relying on manual data entry or limited to academic training environments,¹⁰ this project demonstrates automated, API-based extraction and harmonization of routine clinical data across independent community practices. The multi-site extraction process, harmonized data model, and governance framework reflect design principles shared by larger EHR-based research networks such as the Patient-Centered Outcomes Research Network (PCORnet)¹⁹ and Observational Health Data Sciences and Informatics (OHDSI),²⁰ adapted to the distinct needs of community-based naturopathic practice. By creating reusable research infrastructure rather than a study-specific dataset, this project establishes a foundation for future investigations across diverse conditions, therapies, and patient populations.

Practice characterization

The assembled database characterizes community-based naturopathic practice at a scale and breadth not previously achievable from survey or administrative data. Patients were predominantly female, concentrated in middle adulthood, and commonly presented with chronic health concerns, findings that are broadly consistent with prior descriptions of naturopathic care-seekers.⁵ The age distribution reflects the diverse practice profiles across participating clinics, including one pediatric/family practice and one clinic primarily serving an older adult population. Extending earlier database efforts conducted primarily within academic training clinics,¹⁰ these observations reflect routine care delivered across multiple independent community practices. Although naturopathic doctors were the predominant provider type, incomplete credential documentation limited detailed workforce characterization. The most prevalent diagnoses — fatigue, vitamin D deficiency, insomnia, hyperlipidemia, and anxiety — reflect a mix of chronic metabolic, lifestyle-related, and behavioral health concerns. Laboratory coverage varied considerably by metric (vital signs 30–48% of adult patients; laboratory values 13–25%), reflecting differences in whether clinics routinely entered externally ordered lab results into CharmHealth’s EHR.

The prominence of nutritional deficiency diagnoses may reflect both naturopathic emphasis on nutritional assessment and clinic-specific documentation practices within the participating network. The database also quantified utilization of therapies that are infrequently captured in conventional healthcare datasets, including acupuncture/TCM and IV therapy, providing a more comprehensive characterization of community naturopathic practice than has previously been available from survey-based research.²¹ Acupuncture orders likely reflect referrals to licensed acupuncturists or care provided by dual-licensed ND/LAc practitioners within participating clinics, as acupuncture is not within the standard naturopathic scope of practice in Washington State.

Analytic utility demonstration

Treatment order frequency and co-occurrence analyses demonstrated the ability to characterize naturopathic practice patterns at the cohort level, including modality utilization, subcategory detail, and relationships among treatment approaches. These analyses identified a consistent core of nutritional supplements, dietary counseling, botanical medicine, and exercise across patients with elevated HbA1c, consistent with prior survey-based observations that naturopathic care commonly combines lifestyle, nutritional, and botanical interventions.^{5,21} The hierarchical cluster structure of the co-occurrence heatmap provides additional construct validity evidence for the extraction pipeline: the recovered pattern mirrors the naturopathic therapeutic order's least-force principle without any embedded assumption about treatment philosophy.¹⁸ That automated NLP extraction independently recovers this clinically expected structure constitutes a form of fidelity validation — evidence that the pipeline is capturing coherent clinical practice rather than extraction artifacts. The longitudinal analysis demonstrated the database's capacity to support outcome-oriented research using repeated laboratory measurements. Within the diabetes subcohort, mean HbA1c decreased by 0.77 percentage points between first and last available measurements, with directionally similar but smaller changes observed among patients with prediabetes. Confounders precluding causal inference are addressed in the Limitations.

Limitations

Sampling and generalizability. The seven participating clinics constitute a convenience sample of volunteer practices, not a probability sample of U.S. naturopathic medicine. Five of the seven participate in the Institute for Natural Medicine (INM) naturopathic residency network, which may differ from independent community practices in documentation standards and clinical focus. Additionally, all seven clinics are geographically concentrated in Washington State. Race and ethnicity data are largely missing (89.1% null), reflecting a systemic lack of structured demographic recording across the source EHRs. This precludes subgroup analyses and limits generalizability to racially and ethnically diverse populations, who are historically underrepresented among complementary and alternative medicine (CAM) users generally.³

Clinic heterogeneity. Methodological challenges arose from clinic-specific variations across three primary areas. First, regarding site dominance, one clinic contributes 38% of total patients, heavily weighting pooled statistics toward a single site; notably, a mixed-effects sensitivity analysis with clinic as a random intercept yielded nearly identical estimates to the primary paired t-test (see Results). Second, regarding demographic and practice skew, one clinic is a pediatric/family practice (47% of patients under 18), skewing pooled age distributions in analyses not restricted to adults. Furthermore, clinic-specific specialties — including high rates of IV nutrient infusion, physical therapy, and stress/mindfulness at individual sites — substantially inflate pooled frequencies for those modalities. Third, regarding patient overlap, cross-clinic patient linkage was not performed; thus, a patient seen at more than one clinic would appear as distinct individuals. However, the impact of this overlap is likely negligible given the clinics' geographic distribution, naturopathic patients' typical single-practice relationships, and the within-patient nature of the primary HbA1c analyses.

Data and measurement quality. Several limitations inherent to EHR infrastructure affect data quality. First, diagnosis prevalence reflects conditions recorded in the EHR problem list, which may include historical, rule-out, or billing diagnoses; the problem list structure does not distinguish active from resolved conditions. The false-negative rate of automated PHI detection cannot be estimated.

Second, regarding date-shifting artifacts, de-identification applies a per-patient date shift ($\pm 0-364$ days) that was not validated against a human-annotated gold standard. The shift preserves within-patient intervals but precludes use of absolute calendar dates and cross-patient calendar-time, seasonal, and day-of-week analyses. Because each patient's offset is independently random, pooled seasonal signals in lab values (e.g., Vitamin D, HbA1c) are decorrelated across patients – attenuating rather than systematically biasing seasonal estimates – meaning the shifted data cannot support research questions requiring seasonal inference.

Third, no minimum follow-up interval was imposed for the Longitudinal HbA1c Cohort. Sensitivity analyses requiring ≥ 90 and ≥ 180 days produced essentially identical results. Nevertheless, patients with repeated HbA1c measurements are likely more actively engaged in care than the broader population, introducing potential selection bias. Finally, prescription drug records reflect only provider-initiated orders, and drug frequency estimates are limited to conservative lower bounds due to free-text entry without normalization against a standard vocabulary such as RxNorm.

Causal inference limitations. The longitudinal HbA1c analysis was designed as a proof-of-concept demonstration and cannot support causal inferences regarding treatment effectiveness. Three primary sources of confounding preclude causal inference. First, regarding regression to the mean, because the cohort was selected on the basis of elevated baseline HbA1c, statistical regression toward the mean would produce apparent declines independent of any clinical intervention. Second, regarding unmeasured concurrent care, antihyperglycemic medications prescribed by external primary care physicians or specialists are not captured in this naturopathic EHR; therefore, observed HbA1c changes may partly reflect concurrent conventional management. Third, regarding attrition bias, sample sizes decline at extended follow-up intervals (visible in Fig. 1), introducing potential differential dropout where the remaining cohort may differ systematically from patients lost to follow-up. Although the primary paired t-test and mixed-effects sensitivity analysis use each patient's full first-to-last span and are not directly affected by attrition at specific intermediate intervals, patients with longer follow-up likely represent a highly selected, more engaged subset of care-seekers.

Treatment extraction. Clinical note format varied considerably across clinics, and extraction captured clinician-documented orders only, not patient adherence, dose, or duration. Because validation reviewed extracted records rather than a random sample of source notes, missed orders (false negatives) could not be assessed and sensitivity could not be estimated. The primary objective was building high-precision extraction infrastructure rather than maximizing recall, so validation prioritized PPV. Estimating sensitivity would require a manually annotated corpus of complete clinical notes, which was beyond the scope of this pilot project. Reported treatment frequencies should therefore be interpreted as

conservative lower bounds on true ordering rates. Some per-category PPV estimates carry wide confidence intervals, particularly for nutritional supplements and exercise (each 80%, 95% CI: 58–92%).

Two structural gaps further limit extraction completeness. First, supplement orders placed through integrated dispensing platforms (e.g., Fullscript) were not fully captured. Dispensing platform references appearing as inline URL text were extractable, but DocumentReference records containing these URLs were excluded from the pipeline prior to FHIR store import, and the scope of this gap is not quantifiable from the current database. This represents a source of underestimation of supplement ordering frequencies, particularly at clinics where integrated dispensing is the primary ordering mechanism. Second, the PHI redaction pipeline's named-entity recognition for person names may have inadvertently replaced botanical and homeopathic remedy names with placeholder tokens — particularly those with Latin or given-name-like forms (e.g., Belladonna, Ignatia, Echinacea) — underestimating frequencies for those categories.

Future directions

The infrastructure established by this pilot is designed for systematic expansion along both technical and organizational dimensions. Because the pipeline is built on the FHIR R4 standard, it is adaptable to any compatible EHR system, not only CharmHealth. The governance framework developed alongside the technical infrastructure — including reusable data use and data sharing agreement templates and multi-site oversight processes — addresses the organizational and legal barriers that frequently exceed technical challenges in multi-site data-sharing initiatives. This substantially reduces the administrative burden of onboarding additional clinics. Expanding to clinics in underrepresented settings outside Washington State, including rural, solo-practitioner, and specialty-focused practices, is the highest priority for improving generalizability. Additional expansion could include other clinical settings that prioritize whole-person health and patient-centered care, such as integrative medicine or functional medicine practices. Separately, standardized race/ethnicity collection should be incorporated into clinic onboarding protocols to address the demographic data gaps identified in this pilot.

The assembled database also represents an immediate analytic resource regardless of further expansion. HbA1c was used here as a single exemplar; the database contains 18 years of longitudinal observations across seven clinics for a broad range of clinically relevant measures (including thyroid function, vitamin D, lipid panels, inflammatory markers, and vital signs), alongside condition prevalence records, clinician-ordered treatment patterns, and prescription data. A priority for future work is extending these analyses beyond the glycemic cohort to characterize practice patterns and detect longitudinal outcome signals for other conditions commonly encountered in naturopathic practice, such as thyroid disorders, fatigue-related diagnoses, and cardiometabolic risk factors. The database is available to qualified investigators through data use agreements.

Longer-term, mapping the current data model to the Observational Medical Outcomes Partnership (OMOP) Common Data Model would enable participation in OHDSI's global federated network and interoperability with infrastructure such as PCORnet, facilitating comparative analyses at a scale not

achievable with naturopathic data alone.¹⁹ This infrastructure could also serve as a foundation for a practice-based research network (PBRN)¹¹ for naturopathic medicine, supporting continuous learning from real-world clinical data that learning health system frameworks identify as central to evidence-based quality improvement.²²

Conclusion

This pilot demonstrates the feasibility of constructing a privacy-preserving, multi-clinic naturopathic research database from routine EHR data using standardized interoperability frameworks. The resulting resource supports characterization of real-world naturopathic practice, treatment pattern analysis, and longitudinal outcomes tracking, while providing a foundation for future practice-based research. Expansion to additional clinics and practice settings and interoperability with broader observational research networks represent important next steps toward a scalable, interoperable naturopathic and whole-person health research network.

Abbreviations

AES

Advanced Encryption Standard

API

Application Programming Interface

BMI

Body Mass Index

BP

Blood Pressure

CFR

Code of Federal Regulations

CI

Confidence Interval

DPP-4

Dipeptidyl Peptidase-4

EHR

Electronic Health Record

FHIR

Fast Healthcare Interoperability Resources

GCP

Google Cloud Platform

GLP-1

Glucagon-Like Peptide-1

HbA1c

Hemoglobin A1c
HDL-C
High-Density Lipoprotein Cholesterol
HIPAA
Health Insurance Portability and Accountability Act
HL7
Health Level Seven International
hs-CRP
High-Sensitivity C-Reactive Protein
IAM
Identity and Access Management
INM
Institute for Natural Medicine
IQR
Interquartile Range
IRB
Institutional Review Board
IV
Intravenous
LAc
Licensed Acupuncturist
LDL-C
Low-Density Lipoprotein Cholesterol
LOINC
Logical Observation Identifiers Names and Codes
MRN
Medical Record Number
ND
Naturopathic Doctor
NER
Named-Entity Recognition
NLP
Natural Language Processing
NPD
Naturopathic Patient Database
NUNM
National University of Natural Medicine
OHDSI
Observational Health Data Sciences and Informatics
OMOP

Observational Medical Outcomes Partnership
PABAK
Prevalence and Bias Adjusted Kappa
PBRN
Practice-Based Research Network
PCORnet
Patient-Centered Outcomes Research Network
PHI
Protected Health Information
PPV
Positive Predictive Value
SGLT2
Sodium-Glucose Cotransporter-2
SOAP
Subjective, Objective, Assessment, Plan
SSN
Social Security Number
TCM
Traditional Chinese Medicine
TLS
Transport Layer Security
TSH
Thyroid-Stimulating Hormone
URL
Uniform Resource Locator

Declarations

Ethics Approval and Consent to Participate. This study was approved by the Institutional Review Board of the National University of Natural Medicine (NUNM IRB approval #CK021726). All participating clinics executed data sharing agreements prior to data extraction, and all researchers with database access signed data use agreements. Informed consent was waived by the NUNM IRB due to the retrospective nature of the study and the use of de-identified data.

Consent for Publication. Not applicable. This manuscript contains no individual patient data, images, or identifying information.

Availability of Data and Materials. The de-identified research dataset is not publicly available because it constitutes a Limited Dataset under the HIPAA Privacy Rule (45 CFR §164.514(e)), retaining dates and other elements that preclude public release. It is stored in Google BigQuery, and access requests may be submitted to Helfgott Research Institute at NUNM (research@nunm.edu) subject to an approved data

use agreement and institutional IRB determination. The data pipeline, keyword classification rules, and analysis code will be made publicly available upon manuscript acceptance. Project name: ND-EHR Research Database. Project home page: https://github.com/pkalnins/nd_ehr_analysis (private until public release upon acceptance). Archived version: a versioned release with a Zenodo-assigned DOI will be minted at public release. Operating system(s): platform independent. Programming language: Python 3.11+. Other requirements: dependencies specified in the repository's requirements.txt (including google-cloud-bigquery, google-cloud-secret-manager, google-cloud-storage). License: MIT License. Restrictions to use by non-academics: none. Project documentation is available in the interim at the project's public site.²³

Competing Interests. The authors declare that they have no competing interests with respect to the research, authorship, and/or publication of this article.

Funding. This work was supported by the RAND Corporation under an internally funded research project (Subcontract No. SCON-00000738, Award No. HCHCA003CB). The funder had no role in study design, data collection, analysis, or interpretation.

Authors' Contributions. **P.K.:** Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **R.B.:** Writing – review & editing. **S.C.:** Writing – review & editing. **M.S.:** Conceptualization, Project administration, Writing – review & editing. **E.S.:** Writing – review & editing. **C.C.:** Conceptualization, Project administration, Writing – review & editing. All authors read and approved the final manuscript.

Artificial Intelligence Disclosure. Two generative AI tools were used in preparing this manuscript. (1) *Code assistance:* Claude Code (Anthropic; claude.ai/code; accessed 2025–2026) assisted in drafting, debugging, and refining data pipeline, analysis, and validation code. All code was reviewed, tested, and validated by the authors before use; the final version is publicly available (see Data Availability statement). (2) *Manuscript preparation:* Claude Sonnet 4.6 (Anthropic; accessed June 2026) assisted with drafting and editing manuscript prose. The authors reviewed and edited all AI-assisted output and take full responsibility for the content of this publication. No AI tool is listed as an author. No AI-generated images are included.

Acknowledgements. We thank the Institute for Natural Medicine for logistical support of clinic recruitment and coordination, and the Helfgott Research Institute at the National University of Natural Medicine for IRB oversight and project logistics support. We thank CharmHealth for technical support in facilitating API access to clinic data.

Clinical Trial Number. Not applicable.

References

1. Fleming SA, Gutknecht NC. Naturopathy and the primary care practice. *Prim Care*. 2010;37(1):119–36. 10.1016/j.pop.2009.09.002.
2. Simon M, Rothenberg A, Jenab A, et al. *Naturopathic Physicians as Whole Health Specialists: The Future is Whole Person Health Care*. Portland (OR): Institute for Natural Medicine and American Association of Naturopathic Physicians; 2021.
3. Nahin RL, Rhee A, Stussman B. Use of complementary health approaches overall and for pain management by US adults. *JAMA*. 2024;331(7):613–5. 10.1001/jama.2023.26775.
4. Schloss J, McIntyre E, Steel A, et al. Lessons from outside and within: exploring advancements in methodology for naturopathic medicine clinical research. *J Altern Complement Med*. 2019;25(2):135–40. 10.1089/acm.2018.0403.
5. Steel A, Foley H, Van De Bradley R, Lloyd I, Schloss J, et al. Overview of international naturopathic practice and patient characteristics: results from a cross-sectional study in 14 countries. *BMC Complement Med Ther*. 2020;20(1):59. 10.1186/s12906-020-2851-7.
6. Chamberlin SR, Oberg E, Hanes DA, Calabrese C. Naturopathic practice at North American academic institutions: description of 300,483 visits and comparison to conventional primary care. *Integr Med Insights*. 2014;9:7–15. 10.4137/imi.s14124.
7. Bradley R, Kozura E, Buckle H, Kaltunas J, Tais S, Standish LJ. Description of clinical risk factor changes during naturopathic care for type 2 diabetes. *J Altern Complement Med*. 2009;15(6):633–8. 10.1089/acm.2008.0249.
8. Seely D, Szczurko O, Cooley K, et al. Naturopathic medicine for the prevention of cardiovascular disease: a randomized clinical trial. *CMAJ*. 2013;185(9):E409–16. 10.1503/cmaj.120567.
9. Oberg E, Bradley RD, Cooley K, Fritz H, Goldenburg JZ, Seely D, Saxton J, Calabrese C. Estimated effects of whole-system naturopathic medicine in select chronic disease conditions: a systematic review. *Altern Integr Med*. 2015;4:192. 10.4172/2327-5162.1000192.
10. Leung BM, Flower G, Cooley K, et al. Assessing the utility of an electronic research database to capture whole systems practice at two naturopathic outpatient teaching clinics. *J Evid Based Complement Altern Med*. 2013;18(3):170–5. 10.1177/2156587212470139.
11. Mold JW, Peterson KA. Primary care practice-based research networks: working at the interface between research and quality improvement. *Ann Fam Med*. 2005;3(Suppl 1):S12–20. 10.1370/afm.303.
12. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J Am Med Inf Assoc*. 2016;23(5):899–908. 10.1093/jamia/ocv189.
13. Kreimeyer K, Foster M, Pandey A, et al. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *J Biomed Inf*. 2017;73:14–29. 10.1016/j.jbi.2017.07.012.
14. HL7 International. HL7 FHIR Release 4 (v4.0.1). 2019. <http://hl7.org/fhir/R4/>. Accessed 7 Jul 2026.

15. US Department of Health and Human Services, Office for Civil Rights. Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. HHS. 2012. <https://www.hhs.gov/hipaa/for-professionals/special-topics/de-identification/index.html>. Accessed 5 Jun 2026.
16. Google LLC, Cloud Healthcare API. data de-identification. Google Cloud Documentation. 2024. <https://cloud.google.com/healthcare-api/docs/concepts/de-identification>. Accessed 5 Jun 2026.
17. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46(5):423–9. 10.1016/0895-4356(93)90018-V.
18. Zeff JL, Snider P, Myers S, DeGrandpre Z. A hierarchy of healing: The therapeutic order. In: Pizzorno JE, Murray M, editors. *Textbook of Natural Medicine*. 4th ed. St. Louis (MO): Churchill Livingstone; 2013. pp. 18–33.
19. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inf Assoc*. 2014;21(4):578–82. 10.1136/amiajnl-2014-002747.
20. Hripcsak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inf*. 2015;216:574–8. PMID:26262116.
21. Boon HS, Cherkin DC, Erro J, Sherman KJ, Milliman B, Booker J, et al. Practice patterns of naturopathic physicians: results from a random survey of licensed practitioners in two US states. *BMC Complement Altern Med*. 2004;4:14. 10.1186/1472-6882-4-14.
22. Friedman CP, Wong AK, Blumenthal D. Achieving a nationwide learning health system. *Sci Transl Med*. 2010;2(57):57cm29. 10.1126/scitranslmed.3001456.
23. Kalnins P. ND-EHR Dataset: naturopathic EHR research database documentation. 2026. <https://pkalnins.github.io/nd-ehr-dataset/>. Accessed 9 Jul 2026.

Figures

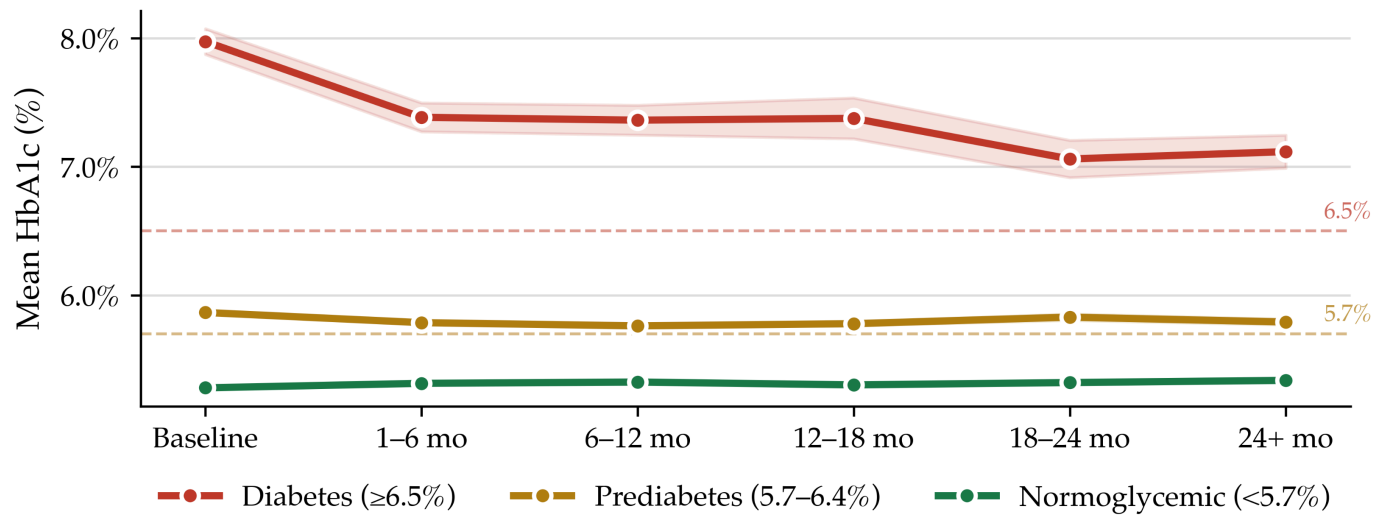


Figure 1

Mean HbA1c trajectory by baseline glycemic group

Figure 1 — HbA1c trajectory

Mean HbA1c (%) at six time intervals from each patient's baseline HbA1c reading, stratified by glycemic category (normoglycemic <5.7%, prediabetes 5.7-6.4%, diabetes ≥6.5%). Shaded bands represent ±1 SE. Sample sizes at each interval are shown below the x-axis. Dashed horizontal lines indicate the prediabetes (5.7%) and diabetes (6.5%) diagnostic thresholds. Longitudinal HbA1c Cohort: n=5,576 patients with ≥2 HbA1c readings on distinct dates.

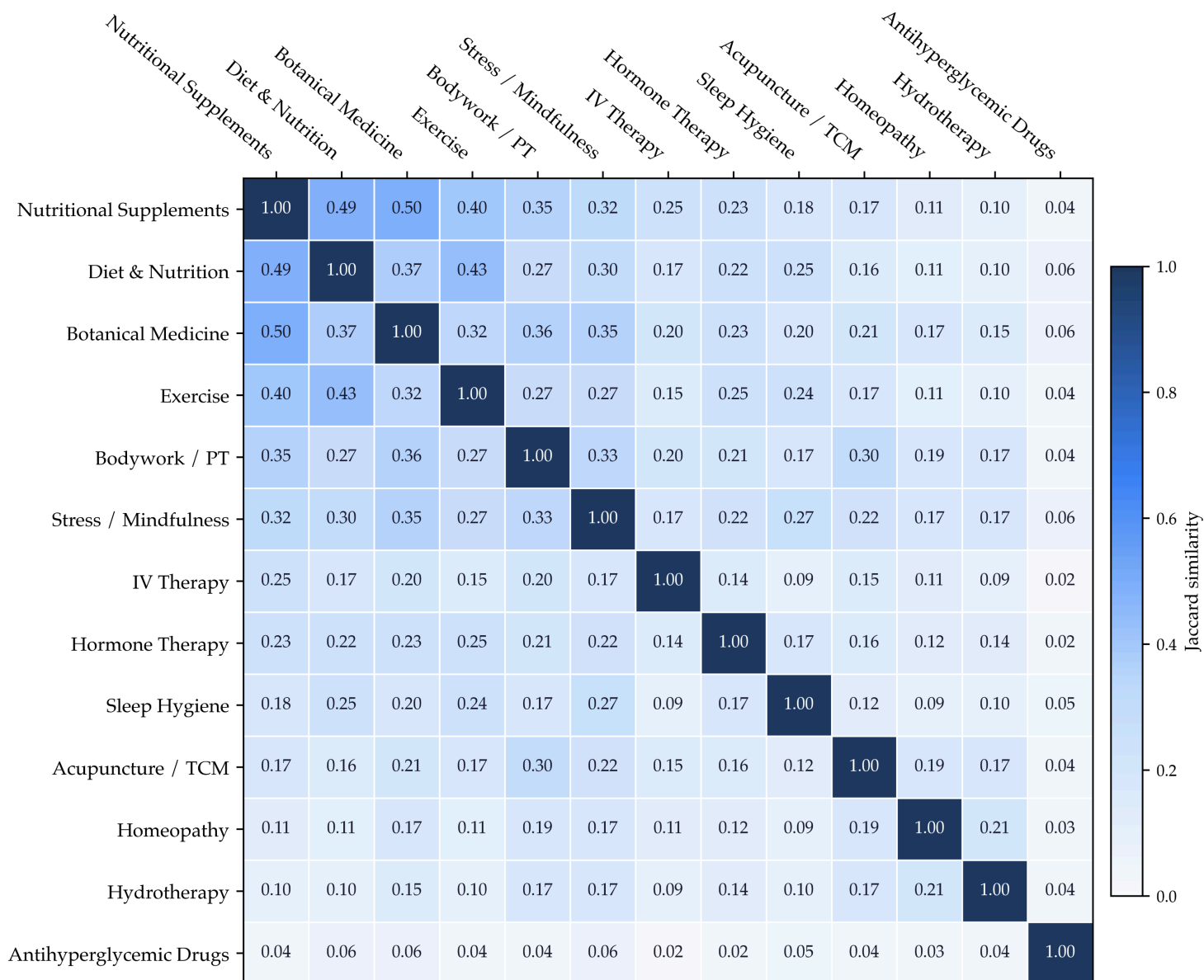


Figure 2

Pairwise treatment co-occurrence — elevated HbA1c cohort

Figure 2 — Treatment co-occurrence heatmap

Pairwise Jaccard similarity matrix for 12 categories of clinician-ordered treatments plus antihyperglycemic drugs in the Elevated HbA1c Cohort (n=2,864). Jaccard similarity = (patients with orders for both treatments) / (patients with orders for either treatment); values range from 0 to 1.0. Diagonal entries are 1.0 by definition. Treatment categories are ordered by prevalence. Hormone therapy combines bioidentical/compounded formulations from clinical notes with prescription hormone records.