Additional file 1: supplementary methods, tables and figures

Authors: Karla Alejandrina Planter-Pérez, MSc; Sofía Gutierrez-Perez, PhD; Judith Carolina De Arcos-Jiménez, PhD; Patricia Noemi Vargas Becerra, PhD; Tadeo Eduardo Hübbe Contreras, PhD; Nancy Evelyn Navarro-Ruiz, PhD; Roberto Miguel Damián-Negrete, MD; Oscar Francisco Fernández-Diaz, MSc; Marco Antonio Romero Güémez, MSc; Elizabeth Bugarin González, MSc; Jaime Briseno-Ramirez, MD (corresponding author)

Contents

Supplementary methods

Method S1. Performance domains and data-source architecture

The available measures were mapped to domains of the WHO Prison Health Framework and the prison-health indicator literature. Environment and resources comprised occupancy, prison medical staffing, general-system medical capacity and reported prison service availability. Access and financial protection comprised treatment receipt, prison-centre provision, unattended care and reliance on personal or family resources. Care processes and prevention comprised examination on arrival, periodic medical review and vaccination, with questionnaire-defined maternal and screening flows. Population-group variation represented equity. Administrative complaints, human-rights requests and externally processed health-related records represented accountability, and DNSP supplied external health-service supervision. Longitudinal mortality indicators comprised repeated state-month death counts in PRS administrative reports and remained distal surveillance signals. "Longitudinal" refers to repeated ecological observations over calendar time, not individual follow-up or prospective data collection; the reported deaths were not clinically adjudicated. Mental-health need and safety were described but not treated as service performance. Clinical quality, disease control, continuity, efficiency, general acceptability, emergency response and clinically adjudicated outcomes were not measured. No composite score or ranking was created.

Eight national public data systems contributed complementary units and estimands. ENPOL supplied individual and person-condition outcomes; ENSANUT supplied independent disease-specific general-population benchmarks; CNSIPEE supplied state prison population, capacity, staffing, service availability and accountability counts; PRS supplied monthly state occupancy and administrative-death reports; DGIS supplied general-system medical capacity; DNSP supplied external centre supervision; CNDHE supplied externally processed health-rights records; and CONAPO supplied marginalisation and population denominators. Parallel components were not treated as one linked cohort or a single causal chain. Table S23 makes the domain–indicator mapping and remaining measurement gaps explicit.

Method S2. ENPOL outcomes, denominators and survey design

The national descriptive denominator comprised chronic person-condition episodes with ongoing care need. Mutually exclusive categories were prison-centre provision, self/family financing, another provider and unattended care. The primary structural outcome was defined at person level as at least one unattended diabetes, hypertension or cancer episode. A two-part episode analysis separated absence of current treatment from provider conditional on treatment. Table S1 gives item-level construction, Table S7 gives the two-part estimands and Table S9 gives the complete sample flow.

Design-based estimates used `FAC_PER`, `EST_DIS` and `FPC`, with the person identifier as the cluster when a respondent contributed repeated condition episodes. Structural association models were model-based because exposures varied at state level. Small-cluster uncertainty was examined with the heteroskedasticity-consistent type 1 (HC1) variance estimator clustered by state, using a `t(31)` reference, parametric bootstrap and state-block bootstrap. Centre-of-interest analyses remained sensitivities because the public centre identifier was available for only part of the analytical sample and no universal validated centre crosswalk exists.

Person-level process indicators were medical examination on arrival (`P1_32`), periodic medical review during imprisonment (`P1_33_1`) and vaccination during imprisonment (`P1_33_2`). Valid yes/no responses were analysed and do not know/no response were missing. National and jurisdiction-specific percentages used the ENPOL design, and federal-minus-state differences were obtained from design-based linear probability contrasts. Holm adjustment covered these three focal contrasts. Questionnaire-defined secondary indicators were centre-provided cervical and breast screening in the previous 12 months, periodic prenatal review and satisfaction with maternal medical care. Screening results were not labelled guideline-eligible coverage, and maternal satisfaction was not treated as general acceptability. Suicidal thoughts and attempts were descriptive need/safety measures, not indicators of service response.

Method S3. ENPOL–ENSANUT cross-survey benchmark

Disease eligibility for cross-survey benchmarking required a comparable physician-diagnosis denominator and a disease-specific current-treatment item or item set in ENPOL 2021 and both ENSANUT rounds. Only diabetes and hypertension met both requirements. Cancer, tuberculosis, hepatitis and HIV lacked an equivalent treatment-status construct across the comparison surveys and remained in ENPOL-only analyses. This compatibility rule was established before the cross-survey contrasts were estimated; selection was not based on prevalence, contrast magnitude or statistical significance.

Diabetes and hypertension were analysed separately among adults aged at least 20 years with a physician-reported diagnosis. Gestational-only diagnoses were excluded from ENSANUT. Absence of current treatment was defined from `P1_26` in ENPOL and from the absence of medication and questionnaire-listed other management in ENSANUT. Each survey retained its own design: ENPOL used `FAC_PER`, `EST_DIS` and `FPC`; ENSANUT used `ponde_f`, `est_sel` and `upm`; isolated primary sampling units used the survey-package adjustment.

Quasibinomial survey models containing age group, sex and their interaction produced marginal standardised estimates; ENSANUT predictions were averaged over the disease-specific ENPOL age-by-sex distribution. Within-survey variances used the delta method. Risk-difference variance combined the independent design variances, and prevalence-ratio uncertainty used the delta method on the log scale with normal-Wald intervals. ENSANUT 2021 was the contemporaneous focal benchmark and ENSANUT 2024 was a secondary repeated population benchmark, not follow-up of ENPOL respondents. Table S11 reports denominator flows and Table S12 reports measurement-definition sensitivities.

Method S4. Structural specificity and equity

DGIS 2021 medical professionals were aggregated by state and divided by CONAPO 2021 population. Establishments identifiable as penitentiary or custodial were excluded from the primary general-capacity measure to reduce construct overlap; an all-facility definition was a sensitivity. Three state-clustered specifications compared prison-specific staffing plus occupancy, general medical capacity plus occupancy and all three exposures jointly. State-block bootstrap used 2,000 replicates. Residualisation was prespecified only if the absolute correlation reached 0.70 or a variance inflation factor (VIF) reached 5.

Equity analyses used quasibinomial design-based models to estimate marginal standardised risks of unattended care, prison-centre provision and self/family financing by sex, age group and Indigenous identification. Contrasts were adjusted for disease and the other demographic variables; adjusted risk differences and prevalence ratios used delta-method

normal-Wald intervals. Holm adjustment covered nine focal contrasts. Inference was suppressed when either group had fewer than 30 unweighted episodes or a Kish effective sample size below 20. Group differences were not automatically labelled inequities because disease mix, selection, survival, centre segregation and custodial conditions may differ.

Method S5. CNSIPEE retrospective state-year panel

Year-specific dictionaries harmonised reference years 2021–2024. The state-year service outcomes were the proportions of reporting centres offering any medical service, medication supply and chronic-disease treatment. Official nonnumeric missing codes, structural absence and non-reporting were preserved. For accountability variables only, a blank child count was recoded to zero when the parent explicitly established absence of the complaint or request (`parent=2`); children under unknown or missing parents remained missing. A state-year service outcome required determinate reporting from at least 80% of observed centres.

Occupancy and medical personnel per 1,000 people deprived of liberty were modelled separately using linear models of centre proportions with state and year fixed effects, state-clustered HC1 variance and a `t(G-1)` reference; coefficients were expressed as percentage-point differences. Holm correction covered six exposure-outcome contrasts. Linear first-difference models with year indicators, complete-reporting restrictions and leave-one-state-out fits were sensitivities. CNSIPEE 2026 reference-year 2025 tabulations were used only as a descriptive bridge because comparable medication and chronic-care outcomes were unavailable.

Administrative medical/psychological complaint records and requests to public human-rights bodies that identified health as an allegedly violated right were analysed as exploratory accountability counts. State-years required at least 80% determinate centre reporting. Poisson quasi-maximum-likelihood estimation (QMLE) models used state and year fixed effects, log prison population as offset, state-clustered HC1 variance and a `t(31)` reference. Holm adjustment covered four exposure–signal models. Negative-binomial fits and leave-one-state-out models evaluated overdispersion and influence. Because reporting access and institutional recording affect the counts, neither signal was interpreted as a monotonic quality score.

Method S6. PRS retrospective state-month panel

The historical panel used the official broad administrative-death category through June 2025. From July 2025, the distinct source category for deaths from natural causes or disease was analysed as a separate short series and was not concatenated with the historical outcome. Three June–August 2019 files contained no statistical tables and were not imputed.

The focal Poisson model included state and calendar-month fixed effects, log end-of-month prison population as offset and one-month-lagged occupancy per 10 percentage points. Inference used state-clustered HC1 variance with `t(31)` and a 2,000-replicate state-block bootstrap. The standardised contrast between 100% and 120% occupancy averaged predictions over the observed state, calendar-month and population distribution and used delta-method uncertainty with the same `t(31)` reference. COVID-period exclusion, state trends, alternative lags, negative-binomial modelling and leave-one-state-out fits were protocol-defined sensitivities. Leverage restrictions, quadratic form and the future-occupancy diagnostic were post hoc. These analyses support surveillance rather than prospective follow-up, cause-specific mortality attribution or a causal threshold.

Method S7. Accountability and external supervision

CNDHE 2024 and 2025 open tables (reference years 2023 and 2024) were searched for accepted conciliations and accepted recommendations that simultaneously identified one of seven health-related violation types and the state penitentiary system as the institution responsible. Official `ND` and `NSS` cells remained missing. The extracted values are procedure–violation-type–institution records: one procedure can contribute to multiple violation types, so sums are not unique cases, unique people or incidence rates. Changing state coverage precluded a two-year trend interpretation.

The DNSP 2021 report was parsed centre by centre for explicit health-service findings in the supervised state-prison universe. A centre was classified as explicit deficiency, explicit appropriate attention, or health finding not disclosed. Non-disclosure was never treated as adequacy. National category counts were reported, and the state share of supervised centres with an explicit deficiency was compared concurrently with ENPOL state estimates of prison-centre provision and unattended care using Spearman correlation.

A 10,000-replicate state bootstrap quantified uncertainty but did not propagate ENPOL sampling uncertainty. Because the supervised universe changes across DNSP editions and report formats changed, no longitudinal trend was constructed.

Method S8. Analytical hierarchy and reproducibility

The ENPOL person-level occupancy model remained the overall primary analysis. Component-specific focal estimands were the ENPOL–ENSANUT 2021 benchmark, the joint DGIS model, six Holm-adjusted CNSIPEE service-availability contrasts and the historical PRS lagged-occupancy model. ENSANUT 2024, equity, provider conditional on treatment, CNSIPEE first differences and the recent explicitly labelled PRS series were secondary. Following a WHO domain-gap audit, a dated post hoc amendment was frozen before estimating the final care-process and accountability contrasts. The three ENPOL jurisdiction contrasts formed one Holm family; CNSIPEE accountability models, CNDHE descriptions and DNSP correlations were exploratory. Measurement-definition, leverage, functional-form and future-exposure diagnostics were post hoc.

Primary ENPOL sensitivities required a documented non-treatment barrier; changed occupancy source, transformation and time alignment; reproduced the legacy Indigenous-response coding; added incarceration duration or number of non-communicable conditions; compared person, episode, all-condition and proportional outcomes; applied state-scaled survey weights; changed the association scale; omitted each state in turn; and used state-clustered, parametric-bootstrap and non-parametric state-bootstrap inference. Staffing sensitivities added DNSP and used log staffing. Two-part models were compared with state-intercept and centre-intercept specifications in the centres-of-interest subset. The CNSIPEE and PRS sensitivity sets are described in Methods S5–S6. Measurement-definition, leverage, functional-form and future-exposure diagnostics were post hoc.

Analyses used R 4.5.3 and Python 3.12.13. Executed R packages were `lme4` 2.0.1, `survey` 4.5, `sandwich` 3.1.1, `MASS` 7.3.65, `haven` 2.5.5, `dplyr` 1.2.1, `purrr` 1.2.1, `readr` 2.2.0, `readxl` 1.4.5, `tibble` 3.3.1 and `ggplot2` 4.0.2, with `parallel` from base R 4.5.3. The original ENPOL component is rebuilt with `bash scripts/run_pipeline.sh`. The benchmark/DGIS, CNSIPEE, PRS and domain-expansion components are rebuilt with their module-specific `run_all.sh` commands reported in the reproducibility archive. Validators check source manifests, joins, denominators, model arithmetic and disclosure gates. The public archive is software-only and contains scripts, methodological protocols and configuration dictionaries, but no source or derived datasets, estimates, tables or figures. Users obtain the public-use microdata and administrative source files directly from the official providers.

Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist for cross-sectional components

Item	Recommendation	Location in manuscript
1	Identify the study design in the title or abstract	Abstract—Methods; Methods—Study design, sources and performance domains
2	Explain scientific background and rationale	Background
3	State specific objectives and prespecified hypotheses	Final paragraph of Background
4	Present key elements of study design early	Methods—Study design, sources and performance domains
5	Describe setting, locations and relevant dates	Methods—Study design, sources and performance domains
6	Give eligibility criteria and participant selection	Methods—Study population and measures
7	Define outcomes, exposures, predictors and confounders	Methods—Study population and measures; Comparative, equity and structural analyses
8	Give data sources and measurement methods	Methods—Study design, sources and performance domains; Additional file 1, Methods S1–S7 and Tables S1–S2

Item	Recommendation	Location in manuscript
9	Describe efforts to address bias	Methods—Analytical hierarchy and statistical analysis; Discussion
10	Explain how study size was determined	Methods—Study population and measures; complete enumeration of eligible ENPOL records
11	Explain handling of quantitative variables	Methods—Study population and measures; Comparative, equity and structural analyses
12	Describe statistical methods, confounding control, subgroup analyses and missing data	Methods—Analytical hierarchy and statistical analysis; Additional file 1
13	Report participant numbers at each stage	Results—Study population and comparative benchmark; Tables 1, 3, S7, S9 and S11
14	Give participant characteristics, exposures and missing data	Results—Study population; Table 1
15	Report outcome events or summary measures	Results; Figure 2; Tables 5, S22 and S24
16	Give unadjusted/adjusted estimates, precision and confounders	Results; Tables 2–5; Additional file 1
17	Report other analyses	Results; Additional file 1
18	Summarise key results with reference to objectives	Discussion
19	Discuss limitations and direction/magnitude of potential bias	Discussion
20	Provide cautious overall interpretation	Discussion; Conclusions
21	Discuss generalisability	Discussion
22	State funding source and funder role	Declarations—Funding

Checklist note: This checklist covers the ENPOL, ENSANUT and 2021 structural components. They comprise independent cross-sectional surveys and ecological state exposures, not one individually linked cohort. No participant recruitment occurred within the study team; eligible records were selected programmatically from public-use files.

STROBE checklist for longitudinal components

Item	Recommendation	Location in manuscript
1–3	Identify design, rationale and objectives	Abstract; Background; Methods—Study design, sources and performance domains
4–5	Present design, setting and dates	Methods—Retrospective administrative components
6	Define units and eligibility	Methods—Retrospective administrative components; Tables S16, S19 and S25
7–8	Define outcomes, exposures and sources	Methods; Methods S5–S7; Tables S16, S19 and S25–S26
9	Address bias	Analytical hierarchy; protocol-defined sensitivities; Discussion—limitations

Item	Recommendation	Location in manuscript
10	Explain study size	Complete enumeration of eligible state-years and state-months
11	Explain quantitative-variable handling	Methods—Retrospective administrative components
12	Describe fixed effects, variance, multiplicity and missing data	Methods—Analytical hierarchy and statistical analysis; Tables S16–S21 and S25
13	Report observations at each stage	Results—Longitudinal administrative evidence; Tables 6, S16, S19 and S25
14–15	Report descriptive and outcome data	Results; Tables S18–S20 and S25–S26
16	Give estimates and precision	Table 6; Tables S17, S20 and S25–S26
17	Report sensitivity and post hoc analyses	Tables S17, S20, S21, S25 and S26
18–21	Summarise, limit and interpret	Discussion
22	State funding	Declarations—Funding

Checklist note: The administrative panels are retrospective repeated ecological observations. PRS reports do not follow identified individuals, CNSIPEE availability is not clinical quality, and neither component is described as prospective.

Supplementary tables

Table S1. Cross-source analytical variable dictionary

Source	Source field or published label	Nature	Analytical construct	Meaning or coding
ENPOL 2021	`ID_PER`	Identifier	Person cluster	Anonymous respondent identifier; clusters repeated condition episodes
ENPOL 2021	`EST_DIS`, `FAC_PER`, `FPC`	Survey-design fields	Design stratum, person weight and finite-population correction	Used together for design-based descriptive and care-process estimates
ENPOL 2021	`CEN_INT`, `NOM_INT`	Centre identifiers	Centre-of-interest subset	Available only for a subset; used in centre-restricted sensitivities, not as a universal centre link
ENPOL 2021	`P1_3`, `SEXO`, `P1_15`, `P1_1A`, `FUERO`	Individual covariates	Age, sex, Indigenous identification, incarceration duration and jurisdiction	Age entered per 10 years or in benchmark age groups; Indigenous do-not-know/nonresponse values were missing; jurisdiction distinguished federal from state/municipal prisons
ENPOL 2021	`P1_24_1:P1_24_8`	Binary diagnosis items	Physician-diagnosed conditions	Diabetes, hypertension, cancer, tuberculosis, hepatitis and HIV were retained; acute respiratory conditions were excluded from the ongoing-care denominator
ENPOL 2021	`P1_26_1:P1_26_8`	Condition-specific categorical items	Current treatment	Code 1 indicated current treatment and code 2 indicated no current treatment
ENPOL 2021	`P1_27_1:P1_27_8`	Condition-specific categorical items	Treatment provider	Codes 1–2 denoted person/family financing, code 3 prison-centre provision and code 4 another organisation
ENPOL 2021	`P1_28_1:P1_28_8`	Condition-specific categorical items	Reason for no treatment	Codes 1, 2, 3 and 6 were documented barriers; codes 4–5 denoted completed treatment/no need; codes 7–9 were other or indeterminate reasons
ENPOL 2021	Derived care regime	Nominal outcome	Centre-provided, self/family-financed, other provider or unattended	Mutually exclusive combination of current treatment, provider and reason for non-treatment

Source	Source field or published label	Nature	Analytical construct	Meaning or coding
ENPOL 2021	Derived person outcome	Binary outcome	Any unattended non-communicable condition	At least one unattended diabetes, hypertension or cancer episode
ENPOL 2021	`P1_32`, `P1_33_1`, `P1_33_2`	Binary process items	Examination on arrival, periodic medical review and vaccination	Yes/no responses; do not know/nonresponse were missing
ENPOL 2021	`P1_34_1`, `P1_34_2`	Binary preventive-care items	Cervical and breast screening	Centre-provided screening in the previous 12 months; questionnaire-defined rather than guideline-eligible denominators
ENPOL 2021	`P1_38`, `P1_40`	Maternal-care items	Prenatal review and maternal-care satisfaction	Periodic prenatal review; satisfied/very satisfied versus other valid responses within the maternal-care flow
ENPOL 2021	`P1_29`, `P1_30`	Binary need/safety items	Suicidal thoughts and attempts	Contextual need and safety indicators, not measures of service response
ENSANUT 2021/2024	`ponde_f`, `upm`, `est_sel`	Survey-design fields	Final adult weight, primary sampling unit and selection stratum	Used within each ENSANUT edition; `estrato` was an urbanicity field and was not substituted for the selection stratum
ENSANUT 2021/2024	`edad`, `sexo`	Individual covariates	Age and sex	Restricted to age 20 years or older and categorised for standardisation
ENSANUT 2021/2024	`a0301`, `a0401`	Diagnosis items	Diabetes and hypertension	Physician-reported diagnoses; gestational-only diagnoses were excluded
ENSANUT 2021/2024	`a0307` plus `A0312A:A0312D`	Treatment items	No current diabetes treatment	No insulin/tablets and no listed other current management
ENSANUT 2021/2024	`a0404` plus `A0408A:A0408E`	Treatment items	No current hypertension treatment	No medicine and no listed other current management
CNSIPEE 2022–2025	`sexostt`	Count	State prison population	Year-end population, summed across determinate centres; denominator for occupancy and staffing rates
CNSIPEE 2022–2025	`capaintt` (2021); `esjupltt` (2022–2024)	Count	Installed spaces	Year-end installed capacity; spaces do not imply staffed beds

Source	Source field or published label	Nature	Analytical construct	Meaning or coding
CNSIPEE 2022–2025	Physician sex/specialty cells selected by official labels	Count	Prison medical personnel	Complete physician components summed and expressed per 1,000 people deprived of liberty; excludes psychology, nursing and other technical personnel
CNSIPEE 2022–2025	`rensnp1` (2021); `rnsnni1` (2022–2024)	Binary centre item	Any medical service	Proportion of centres reporting that the service was offered; availability is not receipt or quality
CNSIPEE 2022–2025	`servme11` (2021); `semebr14` (2022–2024)	Binary centre item	Medication supply	Reported availability; not actual dispensing, adherence or stock continuity
CNSIPEE 2023–2025	`semebr6`	Binary centre item	Chronic-disease treatment service	Available for reference years 2022–2024; structurally absent in 2021 and not coded as no service
CNSIPEE 2022–2025	Year-specific parent, total and health-count fields	Counts with parent–child skips	Medical/psychological complaints and health-related human-rights requests	Explicit absence at the parent item defined a structural zero; unknown/nonresponse remained missing
PRS monthly reports	Published state population and installed-spaces columns	Monthly counts	Occupancy	End-of-month population divided by installed spaces; federal centres excluded from state denominators
PRS monthly reports	Historical administrative-death column	Monthly count	Historical mortality indicator	Broad source category through June 2025; not clinically adjudicated
PRS monthly reports	Natural- or disease-death column	Monthly count	Recent mortality indicator	Distinct source category from July 2025; analysed separately from the historical series
DGIS 2021	Published totals for general physicians, specialists and dentists	Counts	General-system medical capacity	State total per 1,000 CONAPO residents after excluding identifiable custodial establishments
DGIS 2021	Facility name and `CLUES`	Text identifiers	Custodial-establishment exclusion	Used only to identify and audit penitentiary or custodial facilities excluded from the primary general-capacity measure

Source	Source field or published label	Nature	Analytical construct	Meaning or coding
DNSP 2021	Centre score sections and explicit health-service phrases	Text classification	External supervision	Centre classified as explicit deficiency, explicit appropriate attention or no disclosed health finding; non-disclosure was not adequacy
CNDHE 2024/2025	`hechovio_e`; state field; `inpres32`/`inspre37`	Categorical filters and counts	Accepted health-related conciliations/recommendations involving the state penitentiary system	Seven health-related violation types; counts are procedure–violation-type–institution records, not unique cases or people
CONAPO 2020	Published state marginalisation index	Continuous ecological covariate	Marginalisation	State contextual adjustment; not individual socioeconomic status
CONAPO 2021	`POB_TOTAL`	Population count	General-population denominator	State projection aggregated from municipality and sex strata

Footnotes: Field names are retained here solely for reproducibility and are intentionally omitted from the main Methods. Colon denotes a family or contiguous range of source fields, not subtraction. Exact edition-specific CNSIPEE layouts, source files, missing-value rules and model roles are preserved in the archived configuration dictionaries. Source-language labels were translated for exposition; machine-readable field names were not altered. The treatment regime measures reported receipt and provider and does not measure dose, adherence, clinical appropriateness, continuity, stock-outs or clinical outcomes.

Table S2. Data-source roles and linkage constraints

Source	Reference period	Role	Key limitation
ENPOL 2021	2021	Individual outcomes, weights and covariates	Centre identifier populated only for the centres-of-interest subset; stratified by state and centre of interest, simple random sample of persons with no first-stage centre selection
ENSANUT Continua	2021 and 2024	Independently sampled general-population treatment benchmark	Partial endpoint harmonisation; no individual linkage to ENPOL
CNSIPEE editions 2022–2026	Administrative years 2021–2025	State population, capacity, penitentiary workforce and reported centre-service availability	Changing layouts, centre composition, ceiling effects and missing/structural responses
PRS monthly reports	Jan 2015–May 2026	Alternative 2021 occupancy summaries and longitudinal state-month occupancy/death panel	PDF extraction; historical death label is not a clinically adjudicated cause
DGIS Sectoral Health Resources	2021	General-system medical professionals per 1,000 state residents	Ecological capacity measure; excludes custodial establishments in the primary definition
DNSP 2021	2021	Broad penitentiary context	Contains a health-services dimension and overlaps partly with the outcome construct
CNDHE 2024–2025	Reference years 2023–2024	Externally processed health-rights conciliations and recommendations	Counts are procedure–violation-type–institution records; changing state coverage precludes a two-year trend
CONAPO	2020 and 2021	Marginalisation and state population denominator	Ecological measures, not individual socioeconomic information

Footnotes: All 32 state codes linked. The PRS panel contained 384 unique state-month records, and published population minus capacity equalled published absolute overpopulation in all 384 records. Design-based descriptive estimates used the public design stratum, person weight and finite-population correction. ENPOL 2021 is stratified by state and centre of interest; persons were selected by independent simple random sampling within each state stratum from nominal prison-population lists, with allocation proportional to centre size. The design has no first-stage centre selection, so the public design variables fully represent it. For episode analyses, the person identifier was used as the cluster to account for repeated condition episodes. Multilevel structural estimates are model-based and do not incorporate the survey design; small-cluster inference for the 32 state exposure units was therefore probed with state-clustered HC1 variance using a t reference with 31 degrees of freedom, parametric bootstrap and non-parametric state cluster bootstrap (Table S3).

Table S3. Primary occupancy sensitivity analyses

Specification	aOR per state SD	95% CI	P value	n	Inference
CNSIPEE occupancy/capacity, primary person model	1.293	1.103–1.515	0.00351	6,601	GLMM LRT
Strict documented-barrier outcome (`P1_28 in {1,2,3,6}`)	1.331	1.105–1.602	0.00614	6,340	GLMM LRT
`P1_15` do not know/no response coded non-Indigenous (legacy comparison)	1.292	1.103–1.513	0.00350	6,673	GLMM LRT
PRS 2021 monthly mean	1.302	1.121–1.512	0.00170	6,601	GLMM LRT
PRS 2021 monthly median	1.300	1.119–1.512	0.00188	6,601	GLMM LRT
Log occupancy/capacity	1.360	1.146–1.615	0.00132	6,601	GLMM LRT
Previous mixed-year source, state-standardised	1.276	1.085–1.500	0.00596	6,601	GLMM LRT
Person-condition episodes with condition fixed effects	1.270	1.132–1.426	<0.001	8,021 episodes	Logistic, state-clustered HC1
Add number of non-communicable conditions	1.293	1.103–1.515	0.00357	6,601	GLMM LRT
All non-communicable conditions unattended	1.293	1.110–1.505	0.00250	6,601	GLMM LRT
Proportion of non-communicable conditions unattended	1.302	1.115–1.521	0.00229	6,601	Binomial GLMM LRT
State-scaled ENPOL weights, Carle method A	1.312	1.117–1.540	0.00257	6,601	Weighted GLMM LRT
Valid-duration reference sample (`P1_1A`)	1.292	1.102–1.514	0.00361	6,580	GLMM LRT
Add incarceration duration (`P1_1A` categories)	1.280	1.090–1.504	0.00564	6,580	GLMM LRT
Centres-of-interest subset, state random intercept	1.194	0.994–1.433	0.0738	3,920	GLMM LRT
Centres-of-interest subset, centre random intercept	1.208	1.047–1.393	0.0131	3,920	GLMM LRT
Centres-of-interest subset, state and centre intercepts	1.208	1.047–1.393	0.0228	3,920	GLMM LRT, singular
Person-level GLM, state-clustered HC1 variance	1.256	1.112–1.418	0.000589	6,601	Robust t(31)
State cluster bootstrap, states resampled	1.293	1.040–1.688	—	6,601	Percentile, B=2,000

Footnotes: All person models adjust for age, sex, Indigenous identification, DNSP score, state marginalisation and log prison-system size, with a state random intercept unless stated otherwise. In the primary analysis, `P1_15` do-not-know and no-response values were missing; the legacy comparison shows the previous coding. All 2,000 parametric-bootstrap simulations converged; the percentile 95% interval was 1.106–1.521 (seed `20260711`). Across 32 leave-one-state-out refits, adjusted ORs ranged from 1.247 to 1.383 and no 95% confidence interval included one. The duration rows restrict to the 6,580 people with a valid `P1_1A` response and add its five ordinal categories as a factor. Centre-of-interest rows restrict to the 3,920 people in the 44 identified centres across 25 states (Table S10); the nested state-and-centre fit was singular, with state variance estimated at zero and centre variance 0.209. The state cluster bootstrap resampled the 32 states with replacement and refit the model 2,000 times (seed `20260718`); all replicates converged.

Table S4. CNSIPEE supply indicators

State-level indicator, per SD	aOR	95% CI	LRT p	BH q
Medical staff per 1,000 PDL	0.766	0.670–0.876	0.00028	0.0020
Psychology staff per 1,000 PDL	0.868	0.731–1.032	0.116	0.162
Social-work staff per 1,000 PDL	0.877	0.724–1.062	0.181	0.211
Budget per PDL, states with at least 90% denominator coverage	0.817	0.684–0.976	0.0349	0.122
PDL in centres reporting mental-health service	0.869	0.748–1.009	0.0743	0.149
PDL in centres reporting nutrition service	0.875	0.756–1.014	0.0851	0.149
PDL in centres reporting medication supply	0.936	0.787–1.112	0.455	0.455

Footnotes: Each indicator was added separately to a model adjusting for occupancy, state marginalisation, log system size, age, sex and Indigenous identification. Benjamini-Hochberg q values cover the seven exploratory resource indicators. Only medical staff density was prespecified for focused interpretation. Budget coverage retained 23 states and does not survive multiplicity correction.

Table S5. Medical staffing sensitivity analyses

Specification	aOR per state SD	95% CI	LRT p
Main staffing model, excluding DNSP	0.766	0.670–0.876	0.00028
Add DNSP score	0.772	0.676–0.882	0.00039
Log-transformed staffing plus DNSP	0.780	0.690–0.882	0.00028
Exclude state reporting zero medical staff	0.779	0.675–0.899	0.00131

Footnotes: DNSP was excluded from the main staffing model because it contains health-service content; the second row tests sensitivity to conditioning on that partially overlapping construct.

Table S6. Multilevel heterogeneity and association measures

Quantity	Estimate
Empty-model state variance	0.217
Empty-model latent ICC	0.062
Empty-model median odds ratio	1.560
Adjusted-model state variance	0.111
Adjusted-model latent ICC	0.032
Adjusted-model median odds ratio	1.375
Adjusted prevalence ratio per occupancy SD	1.177 (95% CI 1.066–1.299)
Marginal risk difference, occupancy Q3 vs Q1	3.0 percentage points
Marginal prevalence ratio, occupancy Q3 vs Q1	1.206

Footnotes: ICC, intraclass correlation coefficient; Q1/Q3, first/third quartile; SD, state-level standard deviation. The adjusted prevalence ratio uses modified Poisson regression with state-clustered HC1 variance and a t reference with 31 degrees of freedom. In an exploratory analysis, the occupancy-by-communicability interaction was detectable with all included conditions (interaction OR 0.787, 95% CI 0.698–0.886; LRT $p<0.001$) but not after excluding hepatitis (OR 0.860, 95% CI 0.684–1.080; LRT $p=0.186$). Condition-specific slopes were estimated from small or heterogeneous strata and were not used to infer biological or organisational mechanisms.

Table S7. Two-part models of treatment access and provider conditional on treatment

A. Survey-weighted distribution

Stage and category	Unweighted episodes, n	Survey-weighted % (95% CI)
Part 1: non-communicable care-need episodes	8,110	100.0
Unattended	8,110 denominator	18.3 (17.3–19.2)
Any current treatment	8,110 denominator	81.7 (80.8–82.7)
Part 2: treated episodes	6,757	100.0
Prison-centre provided	6,757 denominator	67.4 (66.2–68.6)
Self/family financed	6,757 denominator	30.8 (29.6–32.0)
Other provider	6,757 denominator	1.8 (1.4–2.2)

Footnotes: Estimates incorporate ENPOL design strata, person weights, finite-population corrections and clustering of episodes within people. Under the strict definition, which excluded `P1_28 in {7,8,9}` from the denominator, 14.9% (95% CI 14.0–15.8) of 7,789 classifiable episodes had a documented barrier and 85.1% (84.2–86.0) had current treatment. Part 2 is unchanged because it is already restricted to treated episodes.

B. Adjusted associations: primary state-clustered models

Stage and outcome	Exposure, per state SD	aOR (95% CI)	P value	Episodes; people
Part 1: unattended	Occupancy/capacity	1.27 (1.13–1.43)	<0.001	8,021; 6,601
Part 1: unattended	Medical staff per 1,000 PDL	0.79 (0.71–0.88)	<0.001	8,021; 6,601
Part 2: prison-centre provision among treated	Occupancy/capacity	0.63 (0.48–0.83)	0.0015	6,680; 5,463
Part 2: prison-centre provision among treated	Medical staff per 1,000 PDL	1.12 (0.84–1.49)	0.430	6,680; 5,463

Footnotes: These primary two-part models use logistic regression with condition fixed effects, HC1 variance grouped by 32 states and t inference with 31 degrees of freedom. Occupancy models adjust for DNSP score, state marginalisation, log system size, age, sex and Indigenous identification. Staffing models adjust for occupancy and the same covariates except DNSP.

Multilevel sensitivity

Stage and outcome	Exposure, per state SD	State-intercept GLMM aOR (95% CI)	LRT P value
Part 1: unattended	Occupancy/capacity	1.30 (1.11–1.52)	0.0024
Part 1: unattended	Medical staff per 1,000 PDL	0.77 (0.68–0.88)	<0.001

Stage and outcome	Exposure, per state SD	State-intercept GLMM aOR (95% CI)	LRT P value
Part 2: prison-centre provision among treated	Occupancy/capacity	0.61 (0.38–0.97)	0.045
Part 2: prison-centre provision among treated	Medical staff per 1,000 PDL	1.39 (0.88–2.18)	0.164

Footnotes: All four state-intercept fits were nonsingular. A person-plus-state random-intercept model produced a degenerate Hessian because most respondents contributed one episode and was not retained.

C. Adjusted marginal probabilities at state-exposure quartiles

Stage and exposure	Q1 probability	Q3 probability	Q3–Q1 difference, percentage points (95% CI)
Part 1, occupancy/capacity: unattended	13.6%	16.1%	+2.5 (+1.4 to +3.6)
Part 1, medical staffing: unattended	18.2%	14.3%	−3.9 (−5.6 to −2.2)
Part 2, occupancy/capacity: prison-centre provision	76.3%	68.5%	−7.8 (−11.4 to −3.5)
Part 2, medical staffing: prison-centre provision	66.9%	69.8%	+2.9 (−4.7 to +9.4)

Footnotes: Probabilities were standardised over the observed covariate distribution. Confidence intervals were obtained from 2,000 multivariate-normal draws from the state-clustered coefficient covariance matrix. Part 2 is conditional on treatment receipt and should not be interpreted as a total, mediated or causal effect.

Table S8. Reasons for no current treatment among unattended non-communicable episodes

Reported reason (P1_28)	Unweighted n	Survey-weighted % (95% CI)
Prison centre does not provide treatment	717	56.9 (54.0–59.8)
Other reason	301	19.9 (17.6–22.1)
No money	183	13.3 (11.4–15.2)
Uses home remedies	88	5.7 (4.5–6.9)
Family delivery restricted	44	2.6 (1.7–3.5)
Do not know	18	1.4 (0.7–2.1)
No response	2	0.2 (0.0–0.5)

Footnotes: The denominator is 1,353 state-prison diabetes, hypertension or cancer episodes with no current treatment that were retained in the broad unattended category. Percentages incorporate the ENPOL design variables and clustering of episodes within people. The centre non-provision category was the predominant reported reason.

Table S9. Analytical sample flow

Stage	People	Episodes	Exclusions and notes
ENPOL 2021 completed interviews, national	61,449	—	90.9% of 67,584 selected persons; state and federal prisons
State-prison care-need analytical dataset	7,571	9,159	Excludes federal jurisdiction and records without an included, classifiable chronic episode
Non-communicable primary base	6,674	8,110	At least one diabetes, hypertension or cancer episode
Primary complete-case structural sample	6,601	8,021 model episodes	1 missing age; 72 missing Indigenous identification (`P1_15` 8/9)
Strict documented-barrier model sample	6,340	7,789	Episodes with `P1_28 in {7,8,9}` not classifiable; complete cases

Footnotes: The care-need dataset keeps episodes with current treatment or with no current treatment for a reason other than completed or not needed (`P1_28 not in {4,5}`). Missingness did not justify multiple imputation (0.1% of the non-communicable base for age and 1.1% for Indigenous identification); complete-case coding and the legacy non-Indigenous coding gave effectively identical estimates (Table S3).

Table S10. ENPOL centres of interest to CNSIPEE 2022 crosswalk

Match status	Centres	Primary-sample people
Exact normalised name, unique within state	7	574
Core name after removing institution-type tokens, unique within state	16	1,438
Ambiguous, multiple CNSIPEE candidates within state	2	115
No automatic match; manual adjudication required	19	1,793
Total centres of interest in the primary sample	**44**	**3,920**

Footnotes: The ENPOL public-use file identifies 54 centres of interest (`CEN_INT`/`NOM_INT`); 44 appear in the primary sample (3,920 people, 25 states). The crosswalk matched accent-insensitive, upper-case centre names to the 251 centre-level CNSIPEE 2022 infrastructure records within the same state, first on the full normalised name and then on the core name after removing institution-type tokens, requiring a unique candidate. Ambiguous and unmatched centres were not linked; no universal validated crosswalk exists. The full centre-level listing is `resultados/cen_int_cnsipee_crosswalk.csv` in the reproducibility archive.

Table S11. ENPOL–ENSANUT harmonisation and sample flows

Source and disease	Source records	Diagnosed before exclusions	Gestational-only excluded	Valid treatment and design fields
ENPOL 2021, diabetes	61,449	3,408	Not applicable	3,406
ENSANUT 2021, diabetes	13,402	1,610	19	1,590
ENSANUT 2024, diabetes	12,924	1,633	20	1,613
ENPOL 2021, hypertension	61,449	6,579	Not applicable	6,575
ENSANUT 2021, hypertension	13,402	2,451	59	2,388
ENSANUT 2024, hypertension	12,924	2,570	57	2,513

Footnotes: Diabetes and hypertension were the only ENPOL chronic conditions with comparable physician-diagnosis denominators and disease-specific current-treatment constructs in both ENSANUT rounds; cancer, tuberculosis, hepatitis and HIV remained source-specific. The benchmark analysed each survey independently. ENPOL used `FAC_PER`, `EST_DIS` and `FPC`; ENSANUT used `ponde_f`, `est_sel` and `upm`. Eligible participants were adults aged at least 20 years with a physician-reported diagnosis. Gestational-only diagnoses were excluded from ENSANUT. Absence of current treatment was `P1_26=2` in ENPOL and absence of both medication and other reported management in ENSANUT (`A0307/A0312` for diabetes; `A0404/A0408` for hypertension). Direct standardisation used disease-specific ENPOL age groups (20–39, 40–59 and at least 60 years) and sex. The surveys were not pooled and their weights were not placed on a common scale. Standardised variances came from each survey design; risk-difference variance was the sum of independent survey variances and prevalence-ratio uncertainty used the delta method on the log scale.

Table S12. Cross-survey estimates and measurement sensitivities

Disease/year	ENPOL untreated %	Standardised ENSANUT untreated %	Difference, points (95% CI)	Prevalence ratio (95% CI)
Diabetes/2021	19.91	8.36	11.55 (7.46–15.65)	2.38 (1.50–3.79)
Hypertension/2021	31.73	18.13	13.60 (9.30–17.90)	1.75 (1.39–2.20)
Diabetes/2024	19.91	8.00	11.91 (7.77–16.06)	2.49 (1.52–4.07)
Hypertension/2024	31.73	25.86	5.87 (-1.17 to 12.90)	1.23 (0.94–1.61)

ENSANUT measurement sensitivity	2021 standardised % (95% CI)	2024 standardised % (95% CI)
Diabetes, no pharmacotherapy	17.63 (12.79–22.46)	13.53 (9.39–17.67)
Diabetes, no medication/other treatment excluding fruit-only response	10.34 (6.31–14.38)	8.91 (4.99–12.83)
Hypertension, no pharmacotherapy	25.57 (20.85–30.29)	33.24 (26.66–39.81)
Hypertension, no medication/other treatment excluding fruit-only response	18.13 (14.00–22.25)	25.86 (18.93–32.80)

Footnotes: The endpoint is only partially harmonised. ENSANUT "other treatment" may include diet or exercise and does not demonstrate medication access, clinical adequacy or disease control. Sensitivities were defined after inspection of the questionnaire mapping and are labelled post-specified.

Table S13. DGIS extraction and state join audit

Audit check	Result	Status
DGIS source year	2021	Pass
Source rows / retained after exact duplicate removal	21,662 / 21,661	Pass
Identifiable penitentiary or custodial establishments excluded	24	Pass
Medical professionals excluded with those establishments	264	Pass
DGIS and CONAPO state codes	32 / 32	Pass
Complete DGIS–CONAPO–penitentiary join	32 / 32 states	Pass
CONAPO projected population	128,982,939	Positive
DGIS medical professionals in primary capacity measure	183,642	Positive
ENPOL complete-case denominator after join	6,601 people, 32 states	Pass

Footnotes: The primary general-capacity measure excluded establishments identifiable as penitentiary or custodial to reduce construct overlap with CNSIPEE staffing. An all-facility measure retaining them was a sensitivity. Counts are inputs to an ecological state measure, not evidence of provider availability to an individual respondent.

Table S14. Structural-specificity models and diagnostics

Model and focal term	OR (cluster-HC1 95% CI)	Bootstrap 95% CI	Marginal risk difference, points (HC1 95% CI)
Prison-specific: prison staffing	0.786 (0.701–0.881)	0.669–0.900	-3.32 (-4.84 to -1.80)
Prison-specific: occupancy	1.254 (1.145–1.373)	1.071–1.499	+3.07 (1.96–4.19)
General-capacity: general medical capacity	1.082 (1.018–1.150)	0.796–1.173	+1.12 (0.27–1.98)
General-capacity: occupancy	1.308 (1.190–1.437)	1.111–1.673	+3.63 (2.36–4.90)
Joint: prison staffing	0.797 (0.703–0.904)	0.671–0.919	-3.12 (-4.80 to -1.45)
Joint: general medical capacity	1.039 (0.969–1.115)	0.743–1.125	+0.54 (-0.44 to 1.53)
Joint: occupancy	1.276 (1.160–1.404)	1.061–1.523	+3.30 (2.10–4.50)

Footnotes: All rows use 6,601 people in 32 states and exposures scaled by one state SD. The contrast fixes the focal exposure at -0.5 and +0.5 SD while averaging over observed covariates. State-block bootstrap requested and completed 2,000 replicates. Prison staffing and general capacity correlated at -0.146; VIFs were 1.26 and 1.16. Neither the protocol's `|r|` at least 0.70 nor `VIF` at least 5 residualisation trigger was reached.

Table S15. Standardised equity contrasts

Outcome and contrast	Risk difference, points (95% CI)	Prevalence ratio (95% CI)	Holm p
Unattended, women vs men	-6.67 (-7.80 to -5.55)	0.649 (0.609–0.692)	<0.001
Unattended, 60+ vs 20–39 years	-21.04 (-24.04 to -18.05)	0.328 (0.272–0.396)	<0.001
Unattended, Indigenous vs non-Indigenous	+4.25 (1.83–6.67)	1.246 (1.107–1.403)	0.0026
Centre provision, women vs men	+6.98 (5.59–8.37)	1.128 (1.100–1.157)	<0.001
Centre provision, 60+ vs 20–39 years	+18.73 (14.80–22.66)	1.429 (1.325–1.541)	<0.001
Centre provision, Indigenous vs non-Indigenous	-5.23 (-8.19 to -2.27)	0.907 (0.857–0.960)	0.0026
Self/family financing, women vs men	+0.48 (-0.83 to 1.79)	1.019 (0.968–1.073)	1.000
Self/family financing, 60+ vs 20–39 years	+1.39 (-2.11 to 4.88)	1.058 (0.918–1.219)	1.000
Self/family financing, Indigenous vs non-Indigenous	+0.97 (-1.63 to 3.57)	1.039 (0.939–1.150)	1.000

Footnotes: Risks were standardised over the observed covariate distribution with survey weights and were not reported when either cell had unweighted n<30 or Kish effective n<20; no focal contrast met a suppression rule. Holm adjustment covered the nine contrasts. Sex- and age-group differences are not automatically interpreted as inequities because disease mix, selection, survival, centre segregation and conditions of custody may differ.

Table S16. CNSIPEE longitudinal construction and reporting

Element	Locked rule or validation result
Unit and period	State-year; 32 states × reference years 2021–2024 (128 possible rows)
Outcomes	Share of reporting centres offering any medical service, medication supply and chronic-disease treatment
Missingness	Official nonnumeric missing codes, parent-child structural absence and non-reporting preserved; never recoded as zero
Reporting gate	State-year outcome retained only when at least 80% of observed centres had determinate responses
Exposures	Occupancy per 10 percentage points and medical personnel per 1,000 PDL, modelled separately
Main model	State and year fixed effects; state-clustered HC1; t(G-1)
Multiplicity	Holm correction across six exposure-outcome contrasts
QA	24 staff-subtotal reconciliations, zero mismatches; Oaxaca 2023–2024 staffing missing preserved; Puebla reporting exclusions preserved
2025	Descriptive bridge only; excluded from fixed-effects and first-difference models

Footnotes: The official nonnumeric source codes `NSS`, `ND` and `NP` were preserved as missing and were never interpreted as zero. Changing source layouts and centre composition were mapped with year-specific dictionaries. Chronic-care availability was comparable for three reference years; the other outcomes were comparable for four.

Table S17. CNSIPEE longitudinal models and sensitivities

Exposure and outcome	n state-years	Change, percentage points (95% CI)	Raw p	Holm p
Occupancy, any medical service	124	-0.05 (-0.23 to 0.13)	0.570	1.000
Occupancy, medication supply	124	-1.61 (-4.06 to 0.83)	0.188	1.000
Occupancy, chronic care	92	+0.03 (-3.71 to 3.77)	0.986	1.000
Staffing, any medical service	122	+0.00 (-0.07 to 0.07)	0.972	1.000
Staffing, medication supply	122	+0.18 (-0.91 to 1.26)	0.742	1.000
Staffing, chronic care	90	+0.51 (-2.21 to 3.23)	0.704	1.000

Footnotes: First-difference models were also inconclusive: estimates ranged from -1.11 to +0.91 percentage points across the six contrasts and all intervals included zero. Complete-reporting restrictions reproduced the focal coefficients because retained observations already satisfied the determinate-item rule. All 192 leave-one-state-out fits completed. Short panels, ceiling effects and outcome missingness limit power; the findings do not demonstrate absence of association.

Table S18. CNSIPEE 2025 descriptive bridge

Reference year	States	Prison population	Installed spaces	Aggregate occupancy ratio	States reporting medical staff	Medical staff	Medical consulting-room reporting
2025	32	211,025	176,970	1.192	31	1,281	286/299 determinate centres (95.7%)

Footnotes: Mexico City medical staffing was unavailable under an official nonnumeric missing code; budget was reported for 29 states. Consulting-room reporting was determinate for 286 of 299 centres. The bridge confirms that current national structural tabulations exist, but comparable medication/chronic-service outcomes required for the panel were not available and no 2025 inferential model was fitted.

Table S19. PRS source and outcome-label reconciliation

Audit element	Result
Official monthly attachments expected/found	137/137
Attachments with analyzable statistical tables	134
Cover-only files not imputed	June–August 2019
State occupancy and death cells per analyzable report	32 and 32
Death tables reconciling to source total	133/134
Official discrepancy preserved	August 2017: published 27 vs summed 26
State death cells cross-checked against seven official CSVs	224
2021 state-month occupancy rows reproducing the original extract	384/384

Footnotes: Through June 2025 the outcome used the source's broad administrative-death category. From July 2025, the distinct category limited to deaths from natural causes or disease was analysed as a separate series; the two definitions were not concatenated as if equivalent. Exact source-language labels are retained in the reproducibility dictionaries rather than used as reader-facing analytical labels. These longitudinal mortality indicators are repeated state-month administrative measures, not clinically adjudicated cause-specific deaths or individual prospective follow-up.

Table S20. PRS focal model and protocol-defined sensitivities

Analysis	State-months	Events	IRR per 10-point occupancy increase (95% CI)
Lagged occupancy, state and calendar-month FE	3,872	5,877	1.087 (1.009–1.171)
Exclude March 2020–December 2021	3,168	4,357	1.076 (0.997–1.161)
State-specific linear trends	3,872	5,877	1.094 (1.034–1.157)
Three-month lagged moving average	3,744	5,753	1.096 (1.012–1.187)
Contemporaneous occupancy	3,936	5,927	1.082 (1.005–1.165)
Complete monthly death reconciliation	3,840	5,850	1.088 (1.010–1.172)
Negative-binomial FE sensitivity	3,872	5,877	1.084 (1.014–1.159)
Explicit natural/disease-death label, 11 months	352	980	1.068 (0.977–1.167)

Footnotes: The focal Poisson model used log end-of-month prison population as offset. State-block bootstrap (2,000/2,000 replicates) gave 1.014–1.177. Predicted rates at 100% and 120% occupancy were 0.214 and 0.252 administrative deaths per 1,000 PDL-month, a difference of 0.039 (95% CI 0.003–0.074). A binary above-100% specification was inconclusive (IRR 1.055, 95% CI 0.789–1.411), so no threshold is claimed.

Table S21. PRS post hoc diagnostics

Diagnostic	State-months	IRR or local IRR (95% CI)
Trim observed occupancy p1–p99	3,794	1.085 (1.009–1.166)
Winsorise observed occupancy p1–p99	3,872	1.091 (1.007–1.182)
Exclude occupancy >200%	3,693	1.052 (1.004–1.102)
Exclude occupancy >180%	3,541	1.047 (0.999–1.097)
Exclude occupancy >160%	3,348	1.051 (0.997–1.109)
Midpoint population offset	3,872	1.087 (1.009–1.171)
Lagged occupancy on bidirectional complete sample	3,840	1.086 (1.008–1.170)
Future occupancy on same sample	3,840	1.083 (1.007–1.165)
Quadratic form, local linear IRR at mean occupancy	3,872	1.056 (0.993–1.124)

Footnotes: The quadratic coefficient had p=0.051 with t(31) inference. Leave-one-state-out rate ratios ranged from 1.053 to 1.103; excluding State of Mexico produced the smallest estimate. Persistence after trimming does not resolve the similar future-exposure association. These post hoc checks weaken temporal and causal claims and identify the PRS result as a surveillance signal.

Table S22. Provider and financing distribution by jurisdiction

System	Episodes, unweighted n	Centre-provided, % (95% CI)	Self/family-financed, % (95% CI)	Unattended, % (95% CI)	Other provider, % (95% CI)
National	10,280	53.9 (52.9–55.0)	20.8 (19.9–21.7)	23.1 (22.1–24.0)	2.2 (1.9–2.5)
Federal prisons	1,121	74.6 (70.7–78.5)	5.0 (3.4–6.5)	18.4 (15.1–21.6)	2.1 (1.4–2.8)
State prisons	9,159	51.9 (50.8–53.0)	22.4 (21.4–23.3)	23.5 (22.6–24.5)	2.3 (1.9–2.6)

Footnotes: Percentages and confidence intervals incorporate the ENPOL design stratum, person weight, finite-population correction and clustering of condition episodes within people. The broad unattended category includes documented barriers and other, do-not-know and no-response reasons. The four categories are mutually exclusive and sum to 100% subject to rounding. CI, confidence interval.

Table S23. Performance-domain indicator matrix and measurement gaps

Domain	Indicators observed	Source and period	Principal measurement limit
Environment and resources	Occupancy; prison staffing; general medical capacity; reported service availability	CNSIPEE/PRS/DGIS/DNSP, 2015–2025	Ecological measures; changing administrative and supervised-centre coverage
Access and financial protection	Current treatment; centre provision; self/family financing; unattended care	ENPOL 2021	Self-report; does not establish adequacy or continuity
Care processes and prevention	Arrival examination; periodic review; vaccination; questionnaire-defined screening and maternal care	ENPOL 2021	Content, timing and guideline eligibility are incompletely measured
Equity	Standardised variation by sex, age and Indigenous identification	ENPOL 2021	Contrasts do not by themselves establish inequity mechanisms
Accountability and rights	Administrative complaints; health-rights requests; accepted health-related conciliations/recommendations	CNSIPEE 2021–2025; CNDHE 2023–2024	Reporting access, institutional recording and changing coverage affect counts
External supervision	Explicit health-service deficiency or appropriate-attention findings	DNSP 2021	Non-disclosure is not adequacy; supervised universe changes
Longitudinal mortality indicators	Broad administrative deaths; explicit natural/disease deaths after July 2025	PRS 2015–May 2026	Repeated ecological state-month measures; outcome-label break; no clinical adjudication or individual follow-up
Not measured	Clinical quality and disease control; continuity; efficiency; general acceptability; emergency response; patient-reported outcomes	No adequate national open source identified	Requires stable centre identifiers, clinical audit or prospective centre-level data

Footnotes: Domains were kept separate; no composite score or ranking was computed. Maternal-care experience and mental-health need/safety measures provide limited context but do not establish system-wide acceptability or mental-health-service performance. CNDHE, National Census of State Public Human Rights Bodies; CNSIPEE, National Census of State Penitentiary Systems; DGIS, General Directorate of Health Information; DNSP, National Penitentiary Supervision Assessment; ENPOL, National Survey of the Population Deprived of Liberty; PRS, Prevention and Social Reintegration.

Table S24. ENPOL care-process, prevention, maternal-care and safety indicators

Indicator	Role	Valid national n	National, % (95% CI)	Federal, %	State/municipal, %	Federal minus state, points (95% CI)
Medical examination on arrival	Focal process	60,867	73.36 (73.02–73.71)	78.81	72.93	5.88 (4.57–7.18)
Periodic medical review during imprisonment	Focal process	61,395	41.77 (41.39–42.16)	50.16	41.10	9.06 (7.51–10.62)
Vaccination during imprisonment	Focal prevention	61,419	77.86 (77.55–78.17)	74.81	78.11	-3.30 (-4.72 to -1.88)
Centre-provided cervical screening in previous 12 months	Secondary prevention	11,692	41.01 (40.87–41.15)	36.17	41.33	-5.16 (-5.67 to -4.64)
Centre-provided breast screening in previous 12 months	Secondary prevention	11,691	30.68 (30.54–30.81)	40.48	30.02	10.45 (9.93–10.98)
Periodic prenatal review	Secondary maternal care	1,214	82.40 (81.96–82.84)	76.47	82.48	Suppressed
Satisfied/very satisfied with maternal medical care	Secondary experience	1,001	72.76 (72.28–73.25)	69.23	72.81	Suppressed
Suicidal thoughts since entry	Contextual need	61,407	11.78 (11.50–12.05)	13.73	11.62	2.10 (0.92–3.29)
Suicide attempt among respondents with ideation	Contextual safety	7,663	34.34 (33.14–35.54)	37.29	34.06	3.23 (-1.54 to 8.01)

Footnotes: Estimates incorporate ENPOL strata, person weights and finite-population corrections. The three focal jurisdiction contrasts remained supported after Holm adjustment (all adjusted $P<0.001$). Maternal-care jurisdiction inference was suppressed because federal unweighted denominators were 17 for prenatal review and 13 for satisfaction. Screening denominators follow the questionnaire flow and are not guideline-eligible coverage estimates. Vaccination type and timing were unavailable. Satisfaction applies only to maternal medical care. Suicidality indicators describe need/safety and not service response. CI, confidence interval.

Table S25. CNSIPEE and CNDHE accountability indicators

A. CNSIPEE national records among centres with determinate reporting

Reference year	Determinate centres, medical/psych complaints	Medical/psych records	Records per 100 complaints	Determinate centres, health-rights requests	Health-rights records	Records per 100 requests
2021	250/251	859	25.3	242/251	1,113	34.3
2022	246/248	7,142	35.3	234/248	1,202	39.9
2023	266/266	4,084	18.4	246/266	2,451	41.3
2024	259/261	6,822	22.8	246/261	2,241	42.1

B. CNSIPEE 2025 descriptive bridge

Jurisdiction	Medical/psych records / all complaints (%)	Health-rights records / all requests (%)
Federal	9,576/34,321 (27.9)	1,492/1,899 (78.6)
State	15,861/46,205 (34.3)	1,499/3,067 (48.9)

C. Exploratory within-state fixed-effects models

Signal and exposure	State-years	Poisson QMLE IRR (95% CI)	Negative-binomial IRR (95% CI)	Interpretation
Medical/psych complaints: occupancy per 10 points	125	1.65 (1.20–2.28)	1.46 (1.08–1.97)	Direction robust to count model
Medical/psych complaints: medical staff per 1,000 PDL	123	0.73 (0.59–0.91)	0.95 (0.81–1.11)	Model-dependent
Health-rights requests: occupancy per 10 points	119	1.38 (0.93–2.05)	1.22 (0.97–1.55)	Inconclusive
Health-rights requests: medical staff per 1,000 PDL	117	0.90 (0.75–1.08)	0.98 (0.89–1.08)	Inconclusive

D. CNDHE externally processed health-related records

Reference year	Accepted conciliations	Accepted recommendations	States numeric for all seven health types
2023	25	2	30/32 for both procedures
2024	16	0	30/32 conciliations; 31/32 recommendations

Footnotes: For CNSIPEE, a blank child count was a true zero only when its parent explicitly indicated that no complaint/request existed; official unknown and missing codes remained missing. A complaint or request may have more than one motive or right, so ratios are not mutually exclusive compositions. Poisson QMLE models used state and year fixed effects, prison population as offset, state-clustered HC1 variance and `t(31)` reference; overdispersion ranged from 27 to 173. Holm-adjusted P values for the four Poisson models were 0.013, 0.021, 0.205 and 0.234, respectively. Counts can reflect both underlying problems and access to reporting. CNDHE values sum procedure–violation-type–institution records, not unique cases or people. State of Mexico contributed 16 of 25 conciliation records in 2023 and had a nonnumeric missing value in 2024; therefore, a decline is not inferred. CI, confidence interval; CNDHE, National Census of State Public Human Rights Bodies; CNSIPEE, National Census of State Penitentiary Systems; IRR, incidence rate ratio; PDL, people deprived of liberty; QMLE, quasi-maximum-likelihood estimation.

Table S26. DNSP health-service findings and ENPOL triangulation

A. Supervised state centres, DNSP 2021

Explicit centre-level classification	Centres, n	Percentage of 233 supervised centres
Health-service deficiency	155	66.5
Appropriate health attention	9	3.9
Health finding not disclosed	69	29.6

B. Concurrent state-level correlations

DNSP exposure	ENPOL state outcome	States	Spearman rho	Two-sided P value	State-bootstrap 95% CI
Share of supervised centres with explicit health-service deficiency	Prison-centre provision	32	-0.467	0.007	-0.710 to -0.141
Share of supervised centres with explicit health-service deficiency	Unattended care	32	0.164	0.369	-0.154 to 0.457

Footnotes: Classification used only explicit phrases in each centre section of the official DNSP 2021 report; non-disclosure was not treated as adequacy. Correlations are concurrent ecological triangulation, not estimates of causality or clinical quality. The 10,000-replicate state bootstrap resampled states but did not propagate ENPOL sampling uncertainty. The supervised-centre universe and report format change across editions, so the table is not a longitudinal trend. CI, confidence interval; DNSP, National Penitentiary Supervision Assessment; ENPOL, National Survey of the Population Deprived of Liberty.

Supplementary figures

Figure S1. Two-part adjusted probabilities at the first and third state-exposure quartiles

Adjusted marginal probabilities from the primary episode-level logistic models. Part 1 shows unattended care among non-communicable care-need episodes. Part 2 shows prison-centre provision among treated episodes. Points denote probabilities at the first (Q1) and third (Q3) quartiles of each state exposure; horizontal lines are 95% confidence intervals and annotations give the Q3-minus-Q1 difference in percentage points. Intervals for probabilities, differences and ratios used 2,000 multivariate-normal draws from the robust coefficient distribution. Models include condition fixed effects and the covariates described in Table S7, with HC1 variance clustered by state. The Part 2 contrast is conditional on treatment receipt.

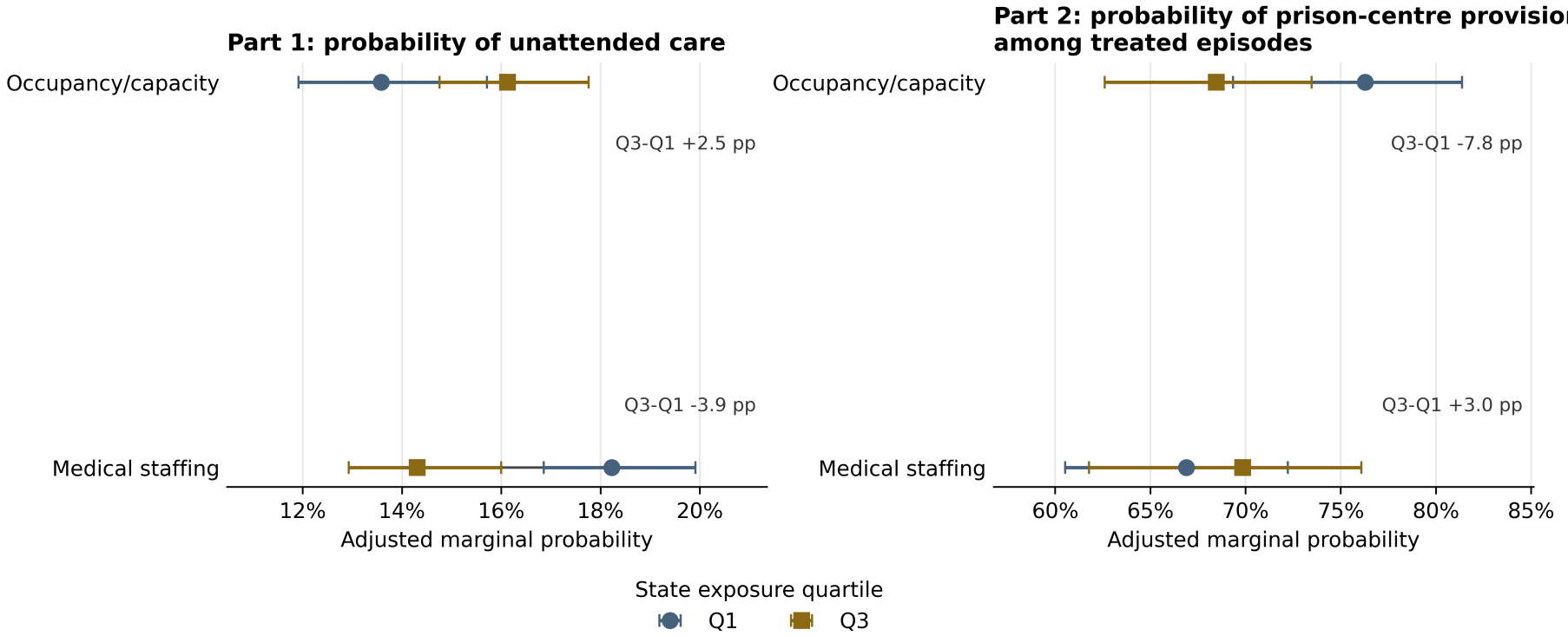

Figure S2. Structural correlates and sensitivity specifications

Adjusted odds ratios and 95% confidence intervals per state-level standard deviation. Diamonds identify the prespecified primary occupancy and staffing exposures; circles identify sensitivity specifications. Occupancy models adjust for age, sex, Indigenous identification, DNSP score, state marginalisation and prison-system size unless an alternative unit or weighting method is named. Staffing models adjust for occupancy, age, sex, Indigenous identification, state marginalisation and prison-system size; DNSP is added only in the indicated sensitivity.