

Source-faithful replication protocol: menopausal hormone therapy and long-term mortality

Bolero

Abstract

Registered research design and conditional interpretation policy; no outcomes observed.

Keywords: Pre-registration, research design

Source-faithful replication protocol: menopausal hormone therapy and long-term mortality

Replication pre-registration: the source paper's results are known; no new replication outcomes have been observed.

Research question

What are the source-defined associations between systemic menopausal hormone therapy and long-term mortality in the original Danish cohort and its specified comparisons?

Scope

Original 1950–1977 female cohort, age 45 entry, all-cause through 2023-07-31 and cause-specific through 2022-12-31, original source analyses only; unresolved inputs/rules remain explicit no-estimate boundaries.

Graph overview

Dependency edges summarize the graph, not unconditional activation. Exact logical conditions and all local outcomes follow below; the separate graph atlas shows each Atom's branches.

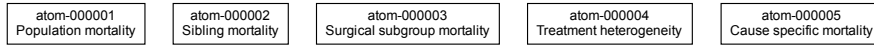


Fig. 1 Registered Atom dependencies

Registered storyline

Reconstruct the published cohort flow, follow-up, uptake and age55 characteristics before describing the nationwide binary and duration-specific all-cause mortality associations. Report the contraindication sensitivity, then the independently interpretable sibling contrast, then cessation sensitivity in the source Table2 order. Describe hysterectomy and BSO age groups with their separately preserved risk sets and death-age summaries; next present treatment form, regimen and initiation age. Finally report cardiovascular, cancer and other mortality at the shorter cause-register horizon. The primary diagnostic and supplementary product-use table remain attached to population mortality; code-only source sensitivities appear in a clearly marked appendix. Every source result is inventoried even if unavailable. No favorable conclusion is prewritten. Each estimate and interval, model provenance and boundary governs its own wording. All modules are canonically active TRUE; their shared data-construction dependencies are checked inside protocols, not manufactured scientific activation edges. Sibling/surgery/exposure/cause evidence informs interpretation of population mortality but does not supply its data or repair an unresolved primary model.

Source verification and search ledger

The source question is fixed. Live searches were limited to source identity, supplement/correction recovery and consequential register definitions. No novelty claim, new hypothesis or redesigned comparison was sought. Original paper and supplement were read in full from supplied bytes; the user-supplied R program was inspected completely without execution. The source hierarchy and access limits appear in replication/source-map.md.

Date/channel	Exact query or request	Result / access level	Use and limit
2026-09-05, live web	"10.1136/bmj-2025-085998" correction supplement; "PMC12915068" menopausal hormone therapy	PMID41708152 PubMed41708152 found; article full text and metadata respectively	Identity corroboration and correction/supplement search; no verified correction located in returned results. Not proof that no correction exists.
2026-09-05 11:55:23 UTC, OpenAlex helper	GET /works/https://doi.org/10.1136/bmj-2025-085998, select id,doi,title,display_name,publication_year,PMID,publisher,date,type,authorships,primary_location	HTTP429; document10.1136/bmj-2025-085998, budget0; no work recorded	Daily budget continued with PMID/publisher/date, type, authorships, primary_location No invented citation neighborhood or saturation claim.
2026-09-05 11:55:23 UTC, NCBI ESearch/ESummary	db=pmc, term="10.1136/bmj-2025-085998", retmax10, sort=relevance; translation="10.1136/bmj-2025-085998" [All Fields]	HTTP200; total1, PMID41708152, summary confirms publication2026-02-085998" [All Fields]	Metadata only from API. Full-text access is independently supplied P/XML, not inferred from ESearch.
2026-09-05, live web	site.bmj.com/content/369/bmj-2025-085998 correction supplement2	Not a peer-reviewed article; authenticated supplement2 or correction returned; unrelated correction-policy results rejected as evidence for this article	Provenance claim; publisher provenance remains unresolved.
2026-09-05, direct BMJ	https://www.bmj.com/content/369/bmj-2025-085998 and /related	HTTP403/301	Cannot independently recover publisher supplemental code or rule out updates. Local P/S retained by hash.

Date/channel	Exact query or request	Result / access level	Use and limit
2026-09-05, live web/official DST	<code>site.dst.dk</code> <code>"VOLUME"</code> <code>"Lægemiddeldatabanken"</code> <code>site.dst.dk</code> <code>"MOR_ID"</code> <code>"FOED_DAG"</code> <code>"BEF"</code>	Official LMDB/BEF single lists read; third-party pages not relied upon	Confirms field labels and catalogue years only. Does not prove VOLUME=DDD or source-export equality.
2026-09-05, source-software lookup	Raw GitHub paths <code>therneau/survival_3.7-0</code>	HTTP404; <code>tested/R/{coxph, cox.zph, tmerge}.R</code> unavailable	No inference from used versioned CRAN archive instead.
2026-09-05 12:12:10 UTC, CRAN archive	<code>https://cran.r-project.org/src/contrib/Archive/survival/survival_3.7-0</code>	<code>R/coxph.R</code> , <code>R/cox.zph.R</code> , <code>R/summary.coxph.R</code> and <code>tmerge</code> source/documentation read without execution	invoked by the supplied source Calls and resolves narrow pre-entry idc semantics, not paper/code scientific conflicts. SHA256 cd96b08ec928b0028f69c942cc788e190b4543c8 Archive remains outside World.

Checked official pages: [PMC article](#), [PubMed record](#), [BEF variables](#), [LMDB variables](#). Full request/error ledgers and search inspection material are retained outside the World. Dates refer to actual retrieval, not article publication. Full source provenance/hash verification is in replication/conversion-audit.md.

P's original background contrasts WHI trials and prior Danish/Swedish observational work; these are retained as contextual positioning with the paper's stated differences in age, treatment and endpoints. No independent reanalysis of those studies, systematic review, current guideline recommendation or literature-saturation claim is made. The original article reports STROBE and restricted register access, but no accessible original SAP/registration identifier was located in the supplied material. This later Bolero registration is not an original-author commitment.

Shared data and construction contract

This is a schema-only, later replication registration of Mikkelsen et al., BMJ 2026;392:e085998 (doi:10.1136/bmj-2025-085998). P denotes the supplied ten-page article PDF; S the four-page supplementary PDF; R the supplied 969-line R script.

These locators refer to the hashes in replication/source-provenance.yaml. The script is read-only evidence, not an executable deliverable. No empirical table or derived author data file was opened.

The unchanged boundary in replication/data-boundary.yaml permits planning with all 531 FSV tables. It asserts no mounted or validated table. All mappings below are candidates until exact extraction provenance is established. Names resembling author files are not sufficient mappings. In particular, FSV TPS is retired civil servants, and KRAF is criminal decisions: neither represents the author's person or cancer file.

Schema evidence and limits

Documentation paths below are relative to `/home/shaoerzhuo/Bolero-denmark/resources/schema/data-regis`. Each cited page's Variables, Coverage and History sections were inspected. Source URL and field evidence are retained in replication/schema-map.csv. Catalogue reference years are not promises of empirical coverage.

Construct / source alias	Exact candidate fields and row/link unit	Documented time and unresolved translation
Person, birth date, sex, maternal/paternal link (<code>t_person2022</code> : pnr, fdato, c_kon, pnrn, pnrnf)	population-and-elections/BEF.md: PNR, FOED_DAG, KOEN, MOR_ID, FAR_ID; person at reference date, REFERENCECID	1985–2026 status snapshots; not the author's full person history. No established mapping for c_status, d_status_hen_start or all historical residents. Need event history and codebook; do not derive a complete cohort from one snapshot.
Immigration/emigration (<code>vnds2022</code>)	population-and-elections/VNDS.md: PNR, HAEND_DATO, INDUD_KODE; many migration events per person	historical register 1973–2024, variable listing in current 2024 extract. Need I/U codes, order, earlier immigration and open departures, not merely first current snapshot.

Construct / source alias	Exact candidate fields and row/link unit	Documented time and unresolved translation
All-cause death (<code>ddato</code> , status 90)	population-and-elections/DOD.md: PNR, DODDATO; death event	1970–2025 historical register, current variable listing 2025. Must verify equivalence to CPR death registration, coverage through 2023-07-31, date corrections and migration-related statuses. Cause-of-death tables are not required for primary mortality.
Birth country (<code>bef2:</code> <code>fland</code>)	BEF.FOEDREG_KODE is birth-registration place; BEF.OPR_LAND is origin, not birth country. FODT.FOED_LAND exists only for the birth-register population (2005–2025).	Neither is automatically <code>fland</code> for women born 1950–1977. Exact historical CPR country mapping is unresolved. No origin/citizenship proxy.
MHT and drug covariates (<code>hc2022</code>)	health/LMDB.md: PNR12, EKSD, ATC, APK, VOLUME, VOLTYPECODE, VOLTYPETXT, NAME, DOSFORM, STRENG, STRNUM, PACKSIZE, VNR; redeemed product transaction	1995–2025. Source paper says prescriptions since 1994, S lists 1995–July 2023. PNR12-to-PNR bridge must be verified. NAME is capital-letter product name; conversion to R’s <code>pname</code> and mixed-case product filters must be established. VOLUME is not documented as DDD by its label. No invented dose conversion.
Diagnoses (<code>lpr77_23_ab</code> , <code>lpr1</code>)	health/LPR_ADM.md: PNR, RECNUM, D_INDDTO, D_UDDTO, C_ADIAG; health/LPR_DIAG.md: RECNUM, C_DIAG, C_DIAGTYPE; contact to many diagnoses via RECNUM	1977–2019; ICD-8 then ICD-10/SKS. Source uses admission date despite discussing discharge codes. No direct PNR on LPR_DIAG. Need source’s merged-row selection, prefix convention and contact deduplication.

Construct / source alias	Exact candidate fields and row/link unit	Documented time and unresolved translation
Modern diagnoses	health/LPR_A_KONTAKT, PNR, DW_EK_KONTAKT, KONT_STARTTIDSPUNKT, KONT_SLUTTIDSPUNKT, ADIAG; health/LPR_A_DIAGNOST, DW_EK_KONTAKT, DIAG_KODE, DIAG_KODE_TYPE	Contacts 2017–2025, diagnoses 2019–2025; transition around 2019. Noted official key validity and original extraction recipe across LPR2/LPR3. Do not append overlapping extracts or equate admissions and modern contacts.
Surgery (op_a1177_23: okode, odto)	health/LPR_SKSOPR.md: RECNUM, C_OPR, D_ODTO (1996–2019); health/LPR_OPR.md: RECNUM, C_OPR (1977–2019; field years have gaps), no operation date listed; link to LPR_ADM	An admission date is not an authorized substitute for missing surgery date. Source recipe for unified okode/odto is absent.
Modern surgery	health/LPR_A_PROCREGISTERING, DW_EK_KONTAKT, DW_EK_PROCEDEREGISTERING, PROC_KODE, PROC_STARTTIDSPUNKT, PROC_SLUTTIDSPUNKT; many procedures per contact	Procedure-to-contact cardinality and source event selection remain unverified.
Historical cancer file	R212–219 requires pnr, search, d_diagnosedato from cancer.csv, joined to LPR diagnoses; S1 states Cancer Register 1943–2017	No verified physical table/field mapping within the supplied FSV catalogue. This is an unavailable mapping, not a declaration that cancer never exists in Danish data. LPR-only exclusion would change the source cohort and is not authorized.

Construct / source alias	Exact candidate fields and row/link unit	Documented time and unresolved translation
Parity, BMI, smoking (<code>gravdb_steen</code>)	health/MFR.md: CPR_MODER, CPR_BARN, FOEDSELSDATO, FOEDSELSLOEBE- NUMMER, PARITET, BMI_MODER, RYGER- STATUS_MODER; mother-child/birth record	MFR 1997–2018; BMI 2003–2018. Source needs pregnancies 1977–September 2023 and <code>gruppe</code> 1/8, <code>slutdato</code> , <code>bmi</code> , <code>ryger</code> . MFRDFOED, LPRM- FRLF/LPRMFRDF and newer MFR satellites are allowed discovery candi- dates, not substitutes: no supplied recipe equates them to <code>gravdb</code> . Multiple births cannot silently become multiple parity increments.
Civil status (<code>civ_bef</code>)	population-and- elections/CIV.md: PNR, CIVST, CIV_VFRA, CIV_VTIL, 1985–2025; BEF.CIVST/CIV_VFRA status records	Source S1 says 1973– 2019; source merged file has no recipe. Handling widowed/registered partnership codes is not recovered by retaining only R's U/G/F/NA without knowing its upstream conversion.
Education (<code>udda2022</code> : <code>uddniv</code> , <code>hf_vfra</code>)	education-and- knowledge/UDDA.md: PNR, HFAUDD, HF_VFRA, HF_KILDE; status 1980–2025; UDDF.md: same plus HF_VTIL, spell history 1974–2025	Need versioned mapping HFAUDD to source <code>uddniv</code> levels (30,60,70,99), and historical highest- attainment history through 2022. HFAUDD itself is not <code>uddniv</code> .
Income (<code>ind9521</code> , <code>ind1</code>)	labour-income-and- wealth/IND.md: PNR, PERINDKIALT_13, ALDER_ULT_INK, reference-year metadata; annual person income 1980–2024	Source 1995–2021; exact <code>ind_aar/year_ind</code> extraction and record scope needed. ALDER_ULT_INK is age at 31 December, not birthday age. Source inflation and quartiles below are retained.

Construct / source alias	Exact candidate fields and row/link unit	Documented time and unresolved translation
Cause of death (dar , dar1)	health/DODSAARS.md: PNR, D_DODSDTO, C_DOD1, C_LISTE_49, 1970–2001. health/DODSAASG.md: PNR, D_DODSDATO, D_STATDATO, C_DODTILGRUNDL_ACME, C_LISTEA, C_LISTE49, 2002–2022	R aliases k_cpr, c_liste_49, and source-file years do not coincide automatically with FSV field names/versions. Need death versus status-date semantics and extraction provenance.
Recent cause-of-death alternative	population-and-elections/DODSAARSAG.md: PNR, DOEDSDATO, DOEDSSTATUSDATO, DOED-SAARSAG_TILGRUNDLIGGENDE, DOED-SAARSAG_GRUPPERING_A_KODE, DOEDSAARSAG_LISTE_49_KODE	2022–2024; possible remapping 2022 revision, not permission to choose it or extend follow-up.

The live official BEF and LMDB Danish variable pages were checked on 2026-09-05. They corroborate field labels, not joins, source export transformations or identifying variation. Data availability remains at documentation/planning level. Full-register counts printed in metadata were not treated as study counts.

Shared execution gate (incorporated into every Atom)

`REQUIRE_SOURCE(target, dependencies)` checks documentary and governed interface prerequisites before any estimation. If a necessary source rule is conflicting or unspecified, emit `STOP_UNRESOLVED_SOURCE`, identify the exact missing decision and affected component, leave all its numerical values absent, and retain the inventory row. If the rule is specified but its input cannot be mapped/accessed with evidence, emit `STOP_UNAVAILABLE_INPUT`. A software crash or timeout is operational **no_result**, not a completed boundary. These are dispositions of an output, not mortality outcomes. Stop only consuming outputs; do not discard independent source-supported outputs. No runtime agent may select an interpretation of a conflict. Source clarification requires a documented amendment before affected execution.

Within governed execution, assert type/encoding of keys and dates, documented join cardinality, source-specified record order, positive nonoverlapping risk intervals, one death per woman, and conservation of person time across lossless splits. These are integrity assertions, not empirical experiments or statistical acceptance thresholds. Unexplained duplicates or missing required keys stop the affected output; do not

silently deduplicate, omit women or impute. Source-specified first/last selections below are different from arbitrary deduplication. No diagnostic may be used to select a favorable analysis.

Source-derived ordered construction

The following is executable pseudocode for the extracted rules. A step marked unresolved does not license choosing either side. Preserve intermediate row provenance inside the governed environment; only disclosure-approved aggregates leave it.

1. **Population (P2–3; R39–77).** Required logical person row: person key, mother/father keys, date of birth, sex, status and status-start date. R sets parent keys missing when a parent identifier occurs more than 30 times in the full imported person file. Preserve this source rule, not a new family-size screen. R obtains first immigration from the first within-person VNDS row only if I. For successive U→I pairs at least 180 days apart, it proposes censoring at departure+180 days, keeps dates from 1995-01-01 and the first qualifying row; it also uses status-derived departure `edato`. Input order and open emigrations are unresolved (D02). P requires women born 1950–1977, alive/resident on the 45th birthday, immigration before age 25. R uses `c_kon K`, excludes `c_status 5/7`, sets `edato` for statuses matching ¹, death date for 90; `start=max(fdato+45365.2, numeric expression as.IDate(1995-01-01), first_ind)`, `stop=min(2023-07-31, edato, cens6_dato, ddato)`, retaining `start<stop`. R includes immigration `<=25365.2`. Calendar birthdays versus 365.2, unquoted date and entry-boundary inclusivity are D01–D02. No exact cohort count or risk set is authorized until these are resolved. Do not replace the numeric expression or silently use 365.25.
2. **Prescription transaction recipe (commented R80–138; active reads R140–141).** Select ATC prefixes G03FA/G03FB/G03CA/G03DA/G03DB/G03DC/G02BA, excluding G03HB from intake. After source-compatible product naming, exclude Prolutex, Vagifem, Estrin, Rewellfem, Oestrin, Premarin, Crinone, Cyclogest, Lutinus, Pleyris, Progestan; exclude vaginal Ovestin. Sum APK in source groups defined by person, date, product name and retained product attributes; verify original `id` grouping and `dcast` keys. P excludes low-dose vaginal estrogen; do not generalize this list into a new exclusion. Decode TABOTR/TAB/KAPB/TABFILM as oral, DEPPLA/GEL/KUTSPRO as transdermal, IUTIND intrauterine, NASSPRO nasal, SUP rectal, INJVSJ injection; source lists VAGCREM/VAGTAB/VAGIND/VAG/VAGGEL as vaginal. Other forms are not reassigned without source evidence.
3. **Episodes (R103–132).** R computes `end=redemption+APK VOLUME with prolong=0`; *G02BA03 intrauterine progestogen lasts 5365.2 days except Jaydess 3365.2. R classifies G03FA continuous combination, G03FB (and unreachable G03HB01 branch) cyclic combination, G03CA estrogen, G03D[A-C]/G02BA03 progestin. Split successive non-progestin records if previous end<current start or product/ATC/strength/person changes; aggregate by split to earliest start, latest end*

¹2-8

and summed *APKVOLUME*. This is an adjacent-record recipe, not an authorized stockpiling algorithm. Ordering, physical volume versus DDD, and accumulated future purchases assigned at episode start are D03/D04. Match estrogen episodes to separate progestin for the same woman when either progestin endpoint lies inside the estrogen episode. R selects the first record ordered by split and descending match indicator. A covering progestin interval with both endpoints outside the estrogen interval is missed; competing matches have no scientific priority. P says simultaneous use. D04 stops affected exposure targets; no repaired overlap join is authorized. Record *htstart/htstop* as rounded $(\text{date-fdate})/365.2$ to 3 decimals only after time-rule resolution.

4. **Clinical history (S1; R174–222).** Extract incident dates using exact code lists in *replication/clinical-codes.yaml*; merge hospital and cancer-register breast/gynaecological events, earliest date per woman. Hysterectomy is first qualifying surgery; bilateral oophorectomy is first bilateral code or second unique unilateral event. R's unilateral accumulation lacks an explicit chronological order and does not establish laterality. D05 records this ambiguity and the absent unified surgery recipe. Thrombophilia/liver/arterial/venous/breast-gynaecological cancer, first MHT, and prior BSO exclude at $\leq \text{start_age}$ in R291–297, while P says before entry (D01). Exclusion counts overlap; count their union separately. Source uses all listed diagnoses; do not add cancers, indications or controls.
5. **Medication covariates (S1; R143–168,193–205).** Diabetes: first A10 except Wegovy Flextouch. Hypercholesterolaemia: first C10AA/C10AB. Hypertension: first use of second distinct source antihypertensive class, using ordered assignment precedence in *clinical-codes.yaml* (a single product is assigned one final class in R152). Heart failure requires qualifying diagnosis and C03C prescription, $\text{date}=\max(\text{first qualifying diagnosis date}, \text{first qualifying loop prescription})$. Atrial fibrillation and valvular disease use first qualifying diagnosis. Whether admission timing implements the paper's discharge wording is retained in D05; no substitute date. Poor general health is R's count of primary diagnosis rows (*C_DIAGTYPE A*) in age [44,45), indicator ≥ 3 . Need source contact-row uniqueness before equating this to three contacts. No extra validation study.
6. **Pregnancies (R224–230,266,311,348).** In verified *gravdb*-equivalent rows, retain *gruppe* 1 or 8, order by *person/slutdato*, assign parity sequentially; smoking values ≥ 1 and < 99 become 1; -1/missing/99 become missing; 0 remains 0. Round BMI to 1 decimal, set $\text{BMI} < 10$ or > 100 missing, carry BMI and smoking forward within woman across pregnancy rows. Baseline parity is latest pregnancy before age45, with absence set to zero in code; subsequent pregnancies increment *cumtdc*; categories [0,1],[1,3],[3,20). *gravdb* provenance/*gruppe* meaning/multiple pregnancy records are D06. These explicit source rules do not prove MFR equivalence or justify treating missing linkage as nulliparity.
7. **Demographics (R58–65,238–255).** Birth country file: unique *pnr/ie_type/fland*, order *person/fland*, first per person, Latin-1 conversion of $\emptyset$ to *oe*. Educational levels: < 99 primary/lower secondary, ≥ 30 upper secondary, ≥ 60 bachelor, ≥ 70 master, 99 unknown (assignments in that order). Keep first record per person/level and source attainment date; source missing date becomes

age0. Civil file: source U/G/F/NA map unmarried/married/divorced/unknown; missing start becomes birth date. No invented widowed mapping. Income: remove source-missing PERINDKIALT_13, multiply by $1.02^{(2023-\text{year_ind})}$, cut at pooled source-file 0/25/50/75/100 percentiles, include.lowest=TRUE, q1-q4. Select latest ALDER_ULT_INK $\leq$ 45 and $\leq$ 55 respectively. Do not compute annual quartiles or choose a different income base. Upstream population/year scope, tie order and repeated quantile breaks require D07. No default rank tie break.

8. **Start-stop data (R258–382).** Left join source one-person incident/baseline tables to person cohort. `tmerge` event at `death_age`; time-dependent indicators for education, civil status, surgery, contraindications, each listed disease, first cessation, cumulative episode starts/stops; cumulative pregnancy count. Split at ages50/55/60/65/70/75 and 2002-01-01,2009-01-01,2016-01-01,2023-01-01 expressed on source age scale. Use survival 3.7-0 counting-process (start,stop] semantics, including its documented same-time tdc/event behavior; do not invent an ordering that differs. `risktime=tstop-tstart`; source calendar start=`decimal_date(entry)+cumulative prior risktime`, year categories [1994,2002],[2002,2009],[2009,2016],[2016,2024). The first label corresponds to P's 1995–2001 because entry is in 1995. Decimal calendar arithmetic can disagree with explicit split dates (D09); do not infer correct calendar truncation merely from a `tmerge` call.
9. **Rolling values (R346–380).** Roll civil, education, latest observed pregnancy BMI/smoking and age55 income onto interval starts using the source keys, with no new backward imputation. Education and income source keys have more columns than the consuming join; duplicate attainment/time records and exact rolling semantics require verification (D07). Keep age45 income until nonmissing age55 update. Baseline disease values were computed in R272–277, but `cpr7 R382` retains the time-dependent variables; confirm the pinned `tmerge` behavior for pre-entry updates rather than assume baseline columns were used. A source-consistent implementation can omit unused BMI/smoking joins for outputs that do not use them only if joins are demonstrably lossless; this does not repair `gravdb` parity.
10. **Exposure on intervals (R384–427).** Join episode DDD on person/interval start=`htstart`. Missing joined dose becomes zero. Source computes `ht_sum=cumsum(ddd)/365.2`, `ht=1` if `ht_sum>0`, otherwise 0; once exposed remain exposed. Duration cuts [0,.001],[.001,1],[1,3],[3,5],[5,10],[10,100); collapsed [0,.001],[.001,5],[5,100). Those upper limits and epsilon are code, not invented categories. If positive sub-.001 duration or ≥ 100 cannot be assigned consistently with printed categories, stop the affected categorical output. Category doses are all booked at episode start in this code, including purchases aggregated later in the episode (D04); do not silently distribute them daily. For route/regimen, maintain cumulative component totals; unique strict maximum defines dominant type. R401/422 carries tied values forward without by-person grouping. No across-person carry or new tie rule is authorized (D10).
11. **Missing demographics (P4; R444–448).** P says impute most frequent category. R hard-codes married, q4, upper secondary and Denmark for missing

civil/income/education/country. Keep unknown categories for descriptive columns except Table1 income is mode-imputed for disclosure. D07 requires establishing imputation population, weighting and tie convention or documenting the original frozen mapping; no recomputation versus hard-coding choice by convenience. No general listwise-deletion policy is added for other covariates.

12. **Models (P3–4; R450 onward).** All `coxph` models use `age` on `Surv(tstart,tstop,event)` and the covariate formulas listed in the consuming Atom. The source R environment is R4.4.1, `survival3.7-0` and `data.table1.16.2` (P9). Preserve the pinned `coxph` defaults where the call omits options: Efron ties, model-based covariance, and `summary(conf.int)` exponentiated coefficient/95% bounds. A source strata term changes baseline hazards; it does not introduce cluster-robust variance. No clustered sandwich, weights, frailty, extra interactions, competing-risk estimator or imputation is added. Source environment/options must be checked, including missing-value behavior, factor ordering and contrast options; unestablished settings with scientific consequences stop that fit. Store coefficient names, not positional recycling from display code. P declares two-sided $P < .05$ for statistical significance; this is not a replication success criterion. No multiplicity correction, power analysis, equivalence margin or new PH acceptance threshold is specified; not added.

Pinned software clarification. The CRAN `survival3.7-0` source archive was retrieved and read without execution (hash and access in `literature.md`). `R/coxph.R` lines3–4 and300–315 verify Efron and nonrobust defaults for these calls without weights, cluster or id arguments; `summary.coxph.R` line1 sets `conf.int=.95`. `R/cox.zph.R` line2 specifies `transform='km'`, `terms=TRUE`, `singledf=FALSE`. Preserve transformed plotting coordinates and original tick labels separately in `schoenfeld_ticks.csv`; do not treat the KM transform itself as age. The archive's `man/tmerge.Rd` explains that `tdc` without `x` starts0 and becomes1, pre-entry updates apply to later intervals, exact touching-boundary events go to the earlier interval and counts to the later, and `same-id/same-time tdc/event` uses last nonmissing value whereas cumulative operators add. This resolves the narrow concern that dropping unused baseline `_b1` columns alone would necessarily lose pre-entry disease information: the source uses `tdc` without `x` on complete incident-age inputs. It does not resolve absent history, duplicate input order or any source scientific conflict.

Formula aliases used by all Atoms. `A_R = year_cat + fland_mode + udda_mode + q_ind_mode + parity_cat + dm + hypertens + afli + valv + hf + statin + many_diag + civst_mode` is the adjustment covariate set in R461. `A_P` is that set without `civst_mode`, corresponding to the printed Table2 footnote. These are names for extracted source specifications, not permission to fit competing versions. D08 governs attribution to published adjusted results. Each Atom states any source-specific removal or addition; such formulas inherit the shared source gates.

Output contract

Each Atom writes governed aggregate outputs under `results/atom-NNNNNN/` at run-time, relative to the governed execution workspace outside the native World (these

paths are interfaces, not present observations). The World keeps exactly atoms/, molecules/ and prereg/. Common `estimates.csv` is long, one row per source target, source exposure row and component:

`target_id, exposure, component, status, reason, source_locator, n, events, person_years, rate_per_10000, estimate, lower, upper, p_value, model_id, unit.`

Components are `summary`, `crude`, `adjusted` only where source has them. Numerical fields irrelevant to a component are blank. `status` is `available`, `reference`, `unresolved_source`, `unavailable_input`, `not_estimable`, `suppressed`, `not_run`, or `operational_failure`. These are interface labels, not new native states. A reference HR=1 is a source-defined comparator only; it is never a replacement for a failed non-reference fit. Unknown CI stays unknown, and an unavailable adjusted estimate does not erase a available crude estimate or denominator. “Available” entails resolved specification, verified inputs, successful fitting and disclosure permission; it does not imply agreement with the publication. Report HR and CI to 2 decimals, incidence rate to 1 decimal, person-years rounded as in source, retaining full precision internally. Preserve P values only from the specified fit; do not infer P from a rounded CI.

`coverage.csv`: `target_id`, `component`, `status`, `reason`, `source_locator`, dependency (logical input or rule), allowing every gap to be traced. `adjudication.md`: ordered local coverage state plus complete component-specific estimates/intervals, mixed patterns, limitations and exact source comparisons. No composite evidence score. `source-comparison.csv`, when authorized results exist, may align source public values with separately labelled rebuilt values and arithmetic difference/printed rounding; no tolerance or success flag.

For scalar/descriptive outputs use `summaries.csv`: `target_id`, `group`, `statistic`, `value`, `unit`, `status`, `reason`. Each requested statistic has its own row and disposition. The source target catalogue fixes requested rows. `flow.csv`, `characteristics.csv`, `products.csv`, and Schoenfeld inputs are additionally defined in `Atom1` and `paper/figures.yaml`. Do not place confidential person keys in the public figure interface. Residual-point disclosure requires governed approval; if unavailable show the reason, not synthetic points.

Identification and scientific scope

The fixed question is the original systemic-MHT mortality question. The paper is a nationwide observational cohort, not random assignment, a new target trial, or a policy discontinuity. The registered primary comparisons relate time-varying ever-use and cumulative exposure to mortality on an age time axis. “Intention-to-treat” is the paper’s label for persistent ever-use after treatment begins; it does not create randomized treatment allocation.

The source’s rationale is that mortality is a severe integrated outcome and that contraindication censoring may be informative for mortality. Its contraindication and cessation analyses change censoring, and the sibling analysis changes the comparison population/baseline hazard. They are retained with their original limitations, not treated as instruments or independent proofs of causal identification. Family strata can address shared familial characteristics but do not remove nonshared treatment

selection. Surgery and treatment strata are source-specific observational comparisons with potential confounding by indication and overlapping membership. Cause-specific Cox estimates are hazards among women remaining at risk under the source’s cause-censoring rule, not cumulative risks in a competing-risk model.

Correct exposure timing, a source-equivalent history, the stated covariate definitions and noninformative censoring assumptions are necessary to interpret the source model. No amount of schema documentation, code agreement or synthetic plotting can establish these empirically. In particular, unresolved calendar, pregnancy, surgery, adjustment-set and source-object conflicts have the stop dispositions in `deviations.md`; a plausible repair would change the question or model and is not authorized.

There is no new causal DAG supplied by the paper. `diagrams/study-dag.dot` is explicitly a source analysis/data-flow illustration; arrows denote input flow, not causal identification. The separate native epistemic graph derives activation and interpretation from Atom/Molecule registrations. Its all-TRUE schedule is fixed; output-dependent wording does not select analyses. Computational input dependencies remain inside protocols. No new graph schema is used.

The paper’s two-sided $P < .05$ convention is preserved only for individual source tests if resolved and performed. Its reported language does not define a replication success/equivalence margin. The package reports estimates and 95% intervals, mixed patterns and missing components. Neither exclusion of 1 nor inclusion of 1 in a CI establishes replication success or equivalence. A source PH plot can constrain interpretation without authorizing new time interactions or a different model. No new power, minimum-event, multiplicity, diagnostic-score or clinical-effect band has been introduced.

Execution, reporting and handoff

This is Stage1 of a two-stage replica lane: source-paper/code conversion, then Transcriber. It has no independent scientific reviewer, author approval, novelty tournament or design debate. The future native accepted/sealed state means frozen/buildable for execution; it does not mean scientific validity or successful replication. This World currently contains pending registrations and a complete results-unfilled manuscript.

The user supplied code after the historical missing-code audit. It was inspected as text only; all 969 lines were read and its hash checked. Its provenance remains user-supplied, not independently verified publisher bytes. Original sources outrank historical aids; unresolved primary-source disagreements are explicit no-estimate targets. No statement of a guessed choice authorizes its implementation. New source clarification must be documented with locator, hash, exact affected outputs and a pre-execution amendment; runtime and Transcriber may not silently repair models or regroup Atoms.

Only schema/documentation and public publication values were accessed. All 531 boundary tables are permitted for planning, with none validated. The registration grants no confidential access or disclosure authority. Governed runtime must establish source-equivalent fields, historical coverage, keys, units and lossless construction

before running each consuming output. Do not contact authors, access their work files, publish, commit, push or release on the strength of this package.

Keep public references under replication/ and current results under future results/atom-NNNNNN interfaces. Every unavailable component keeps its identity, source locator and exact reason; unknown numerics stay blank. Source event-count suppression below3 is preserved for table presentation, but future release still follows actual governed rules. Do not infer a suppressed count from rates or related totals; disclosure review may suppress dependent summaries. Residual-point interfaces may contain sensitive information and need governed clearance; the plot renderer must show unavailable-input explanations if clearance is absent. The initial World contains no such points.

Numerical native design ratings and Molecule C rating exist solely for compatibility and convey no independent review, empirical feasibility, novelty or evidence score. Local boundary/partial/complete states describe execution coverage and never replace the full component evidence. A machine timeout or parse error is operational no_result; a demonstrated unavailable source/input is an evidenced boundary. Current states are not assigned from this source audit.

Transcriber preserves Atom IDs, membership, shared scientific definitions, all replication audit materials, paper.md, results-map.md, figures.yaml and per-figure code. It owns formatting and report builds together with the host. No PDF/LaTeX/full manuscript compilation or ZIP was performed in Stage1. Host-generated README, LaTeX, class/style assets, PDFs and manual views are not a second scientific authority. Synthetic figure fixtures and render tests stay in conversion/source-inspection outside the World; published/source PDFs, screenshots, original uploaded R and extraction caches also stay there or at their original source locators.

Registered interpretation rules

Restrictions take precedence over permissions; residual reporting does not invent a finding.

1. restriction

Condition: TRUE

For every component, unresolved specification, unavailable input, suppression, nonestimation or operational failure prevents a numerical claim for that component. Never copy publication values into replication results, choose a conflicting model, infer null from a boundary, or equate coverage with agreement. Apply this restriction before all permissions.

```
kind: restriction
condition: 'TRUE'
says: For every component, unresolved specification, unavailable input,
↪ suppression,
    nonestimation or operational failure prevents a numerical claim for that
    component. Never copy publication values into replication results, choose
```

a conflicting model, infer null from a boundary, or equate coverage with agreement. Apply this restriction before all permissions.

2. restriction

Condition: `state(atom-000001) == "boundary"`

Population mortality: no source-faithful empirical component is available; state exact affected targets and causes. Other Atoms retain their own scope.

```
kind: restriction
condition: state(atom-000001) == "boundary"
says: 'Population mortality: no source-faithful empirical component is
↪ available;
state exact affected targets and causes. Other Atoms retain their own
↪ scope.'
```

3. primary

Condition: `(state(atom-000001) == "complete") OR (state(atom-000001) == "partial")`

Population mortality: describe each available source-target estimate with its 95% interval, descriptive quantities and availability; retain discordant signs/precision. For partial coverage explicitly list omitted numerical components and why. Do not summarize this coverage state as a scientific verdict.

```
kind: primary
condition: (state(atom-000001) == "complete") OR (state(atom-000001) ==
↪ "partial")
says: 'Population mortality: describe each available source-target estimate
with its 95% interval, descriptive quantities and availability; retain
↪ discordant
signs/precision. For partial coverage explicitly list omitted numerical
components and why. Do not summarize this coverage state as a scientific
verdict.'
```

4. restriction

Condition: `state(atom-000002) == "boundary"`

Sibling mortality: no source-faithful empirical component is available; state exact affected targets and causes. Other Atoms retain their own scope.

```
kind: restriction
condition: state(atom-000002) == "boundary"
says: 'Sibling mortality: no source-faithful empirical component is available;
state exact affected targets and causes. Other Atoms retain their own
↪ scope.'
```

5. supplementary

Condition: `(state(atom-000002) == "complete") OR (state(atom-000002) == "partial")`

Sibling mortality: describe each available source-target estimate with its 95% interval, descriptive quantities and availability; retain discordant signs/precision. For partial coverage explicitly list omitted numerical components and why. Do not summarize this coverage state as a scientific verdict.

```
kind: supplementary
condition: (state(atom-000002) == "complete") OR (state(atom-000002) ==
  ↪ "partial")
says: 'Sibling mortality: describe each available source-target estimate with
  its 95% interval, descriptive quantities and availability; retain discordant
  signs/precision. For partial coverage explicitly list omitted numerical
  components and why. Do not summarize this coverage state as a scientific
  verdict.'
```

6. restriction

Condition: state(atom-000003) == “boundary”

Surgical subgroup mortality: no source-faithful empirical component is available; state exact affected targets and causes. Other Atoms retain their own scope.

```
kind: restriction
condition: state(atom-000003) == "boundary"
says: 'Surgical subgroup mortality: no source-faithful empirical component
  is available; state exact affected targets and causes. Other Atoms retain
  their own scope.'
```

7. supplementary

Condition: (state(atom-000003) == “complete”) OR (state(atom-000003) == “partial”)

Surgical subgroup mortality: describe each available source-target estimate with its 95% interval, descriptive quantities and availability; retain discordant signs/precision. For partial coverage explicitly list omitted numerical components and why. Do not summarize this coverage state as a scientific verdict.

```
kind: supplementary
condition: (state(atom-000003) == "complete") OR (state(atom-000003) ==
  ↪ "partial")
says: 'Surgical subgroup mortality: describe each available source-target
  estimate with its 95% interval, descriptive quantities and availability;
  retain discordant signs/precision. For partial coverage explicitly list
  omitted numerical components and why. Do not summarize this coverage state
  as a scientific verdict.'
```

8. restriction

Condition: state(atom-000004) == “boundary”

Treatment heterogeneity: no source-faithful empirical component is available; state exact affected targets and causes. Other Atoms retain their own scope.

```
kind: restriction
condition: state(atom-000004) == "boundary"
```

says: 'Treatment heterogeneity: no source-faithful empirical component is available; state exact affected targets and causes. Other Atoms retain their own scope.'

9. supplementary

Condition: (state(atom-000004) == "complete") OR (state(atom-000004) == "partial")

Treatment heterogeneity: describe each available source-target estimate with its 95% interval, descriptive quantities and availability; retain discordant signs/precision. For partial coverage explicitly list omitted numerical components and why. Do not summarize this coverage state as a scientific verdict.

```
kind: supplementary
condition: (state(atom-000004) == "complete") OR (state(atom-000004) ==
↪ "partial")
says: 'Treatment heterogeneity: describe each available source-target estimate
with its 95% interval, descriptive quantities and availability; retain
↪ discordant
signs/precision. For partial coverage explicitly list omitted numerical
components and why. Do not summarize this coverage state as a scientific
verdict.'
```

10. restriction

Condition: state(atom-000005) == "boundary"

Cause specific mortality: no source-faithful empirical component is available; state exact affected targets and causes. Other Atoms retain their own scope.

```
kind: restriction
condition: state(atom-000005) == "boundary"
says: 'Cause specific mortality: no source-faithful empirical component is
available; state exact affected targets and causes. Other Atoms retain their
own scope.'
```

11. supplementary

Condition: (state(atom-000005) == "complete") OR (state(atom-000005) == "partial")

Cause specific mortality: describe each available source-target estimate with its 95% interval, descriptive quantities and availability; retain discordant signs/precision. For partial coverage explicitly list omitted numerical components and why. Do not summarize this coverage state as a scientific verdict.

```
kind: supplementary
condition: (state(atom-000005) == "complete") OR (state(atom-000005) ==
↪ "partial")
says: 'Cause specific mortality: describe each available source-target
↪ estimate
with its 95% interval, descriptive quantities and availability; retain
↪ discordant
signs/precision. For partial coverage explicitly list omitted numerical
```

components and why. Do not summarize this coverage state as a scientific verdict.'

12. supplementary

Condition: done(atom-000001) AND done(atom-000002)

Read population and sibling evidence together as different selected populations and comparison structures; report estimates/intervals separately, with no formal difference test, pooled effect or equivalence rule. A family comparison neither proves removal of all confounding nor repairs a blocked population adjusted result.

kind: supplementary
condition: done(atom-000001) AND done(atom-000002)
says: Read population and sibling evidence together as different selected populations and comparison structures; report estimates/intervals
→ separately,
with no formal difference test, pooled effect or equivalence rule. A family comparison neither proves removal of all confounding nor repairs a blocked population adjusted result.

13. supplementary

Condition: done(atom-000003) AND done(atom-000004) AND done(atom-000005)

Read surgery, treatment and cause evidence as source-specific scope qualifications. Overlapping subgroups and distinct cause hazards cannot be averaged into a single score; one available result cannot stand for an unresolved older-BSO or other-cause adjusted target.

kind: supplementary
condition: done(atom-000003) AND done(atom-000004) AND done(atom-000005)
says: Read surgery, treatment and cause evidence as source-specific scope qualifications. Overlapping subgroups and distinct cause hazards cannot be averaged into a single score; one available result cannot stand for an unresolved older-BSO or other-cause adjusted target.

14. residual

Condition: TRUE

All remaining pending, unrun, operationally failed, inconclusive or incompletely classified scope remains explicit. If no evidence exists, state that no replication finding is available. No original source or package completion alone establishes replication agreement.

kind: residual
condition: 'TRUE'
says: All remaining pending, unrun, operationally failed, inconclusive or incompletely classified scope remains explicit. If no evidence exists, state that no replication finding is available. No original source or package completion alone establishes replication agreement.

Preregistered epistemic graph

Each Atom is an interpretable question or diagnostic judgment. Local result states determine registered activation.

atom-000001

What is the source-defined association of systemic MHT with all-cause mortality in the eligible nationwide cohort?

Activation: TRUE

- **boundary:** After a completed evidence-backed assessment, no requested empirical component can be produced because every component is source-underdetermined or unavailable. Operational failure is excluded from this region. — Report precise no-estimate boundaries. There is no empirical mortality finding or replication-agreement conclusion.
- **partial:** At least one requested empirical component is available and at least one is explicitly unresolved, unavailable, not estimable or disclosure-suppressed after completed assessment. No operationally unfinished component is hidden by this state. — Report the available target-specific estimates, intervals and descriptions and all unavailable components. Partial coverage is not statistical confirmation or a uniform evidence direction.
- **complete:** All requested empirical components within this Atom, including code-only targets explicitly assigned to it, have resolved source specifications and completed reportable outputs; no unavailable component is omitted. — Report the entire source-aligned result vector, including mixed estimates and intervals. Complete coverage is not evidence of agreement, significance, equivalence, causality or successful replication.
- **otherwise:** otherwise — State the precise unresolved interpretation/coverage issue and any available component evidence without forcing it into a scientific sign or replication-success category. Unrun/operational failure remains a terminal, not this state.

atom-000002

What mortality association is identified by the source comparison of sisters discordant in eventual MHT use?

Activation: TRUE

- **boundary:** After a completed evidence-backed assessment, no requested empirical component can be produced because every component is source-underdetermined or unavailable. Operational failure is excluded from this region. — Report precise no-estimate boundaries. There is no empirical mortality finding or replication-agreement conclusion.
- **partial:** At least one requested empirical component is available and at least one is explicitly unresolved, unavailable, not estimable or disclosure-suppressed after completed assessment. No operationally unfinished component is hidden by this state. — Report the available target-specific estimates, intervals and descriptions

and all unavailable components. Partial coverage is not statistical confirmation or a uniform evidence direction.

- **complete:** All requested empirical components within this Atom, including code-only targets explicitly assigned to it, have resolved source specifications and completed reportable outputs; no unavailable component is omitted. — Report the entire source-aligned result vector, including mixed estimates and intervals. Complete coverage is not evidence of agreement, significance, equivalence, causality or successful replication.
- **otherwise:** otherwise — State the precise unresolved interpretation/coverage issue and any available component evidence without forcing it into a scientific sign or replication-success category. Unrun/operational failure remains a terminal, not this state.

atom-000003

How does the source mortality association differ in its hysterectomy and bilateral-oophorectomy populations?

Activation: TRUE

- **boundary:** After a completed evidence-backed assessment, no requested empirical component can be produced because every component is source-underdetermined or unavailable. Operational failure is excluded from this region. — Report precise no-estimate boundaries. There is no empirical mortality finding or replication-agreement conclusion.
- **partial:** At least one requested empirical component is available and at least one is explicitly unresolved, unavailable, not estimable or disclosure-suppressed after completed assessment. No operationally unfinished component is hidden by this state. — Report the available target-specific estimates, intervals and descriptions and all unavailable components. Partial coverage is not statistical confirmation or a uniform evidence direction.
- **complete:** All requested empirical components within this Atom, including code-only targets explicitly assigned to it, have resolved source specifications and completed reportable outputs; no unavailable component is omitted. — Report the entire source-aligned result vector, including mixed estimates and intervals. Complete coverage is not evidence of agreement, significance, equivalence, causality or successful replication.
- **otherwise:** otherwise — State the precise unresolved interpretation/coverage issue and any available component evidence without forcing it into a scientific sign or replication-success category. Unrun/operational failure remains a terminal, not this state.

atom-000004

What associations are reported for the source treatment forms, progestogen regimens and initiation-age groups?

Activation: TRUE

- **boundary:** After a completed evidence-backed assessment, no requested empirical component can be produced because every component is source-underdetermined or unavailable. Operational failure is excluded from this region. — Report precise no-estimate boundaries. There is no empirical mortality finding or replication-agreement conclusion.
- **partial:** At least one requested empirical component is available and at least one is explicitly unresolved, unavailable, not estimable or disclosure-suppressed after completed assessment. No operationally unfinished component is hidden by this state. — Report the available target-specific estimates, intervals and descriptions and all unavailable components. Partial coverage is not statistical confirmation or a uniform evidence direction.
- **complete:** All requested empirical components within this Atom, including code-only targets explicitly assigned to it, have resolved source specifications and completed reportable outputs; no unavailable component is omitted. — Report the entire source-aligned result vector, including mixed estimates and intervals. Complete coverage is not evidence of agreement, significance, equivalence, causality or successful replication.
- **otherwise:** otherwise — State the precise unresolved interpretation/coverage issue and any available component evidence without forcing it into a scientific sign or replication-success category. Unrun/operational failure remains a terminal, not this state.

atom-000005

How is the source MHT association distributed across cardiovascular, cancer and other mortality?

Activation: TRUE

- **boundary:** After a completed evidence-backed assessment, no requested empirical component can be produced because every component is source-underdetermined or unavailable. Operational failure is excluded from this region. — Report precise no-estimate boundaries. There is no empirical mortality finding or replication-agreement conclusion.
- **partial:** At least one requested empirical component is available and at least one is explicitly unresolved, unavailable, not estimable or disclosure-suppressed after completed assessment. No operationally unfinished component is hidden by this state. — Report the available target-specific estimates, intervals and descriptions and all unavailable components. Partial coverage is not statistical confirmation or a uniform evidence direction.
- **complete:** All requested empirical components within this Atom, including code-only targets explicitly assigned to it, have resolved source specifications and completed reportable outputs; no unavailable component is omitted. — Report the entire source-aligned result vector, including mixed estimates and intervals. Complete coverage is not evidence of agreement, significance, equivalence, causality or successful replication.

- **otherwise:** otherwise — State the precise unresolved interpretation/coverage issue and any available component evidence without forcing it into a scientific sign or replication-success category. Unrun/operational failure remains a terminal, not this state.

Conversion handoff (no independent scientific review)

- Verdict: `ready_for_transcription`

Paper-to-protocol conversion, not an independent scientific acceptance. Published results are reference targets; no replication has been executed. Native manifest acceptance means the package was compiled and frozen for execution only.

Checks

- `artifact_contract: pass`

Issues

Independent checks

No independent scientific reviewer is part of this two-stage lane. The converter's self-audit is in `docs/replication/conversion-audit.md`.

Converter's source-fidelity self-check

Completed 2026-09-05 by the Stage1 source-faithful protocol extractor. This is a conversion self-check, not independent scientific review, author approval, empirical validation or a finding of replication agreement. The package is prepared for transcription; unresolved scientific specifications remain blocked. No confidential rows or computed register outcomes were accessed, and no author empirical program was executed.

Sources and access

The entire ten-page paper, its supplied four-page supplement, all 969 lines of the supplied R analysis text, the article Markdown/XML derivatives, the supplied intake audit and the historical README were inspected. All seven hashes in the controller's source-provenance.yaml match the supplied files. The provenance and intake audit are copied byte-for-byte into this directory. The full 531-table boundary is also preserved; the schema map records 92 candidate fields across 19 tables. Documentation and schema availability do not establish runtime mounting, historical completeness, valid linkage or identifying variation.

The article PDF governs table layout where the Markdown extraction is malformed. The original Table2/Table3 pages and the supplementary residual plot were also visually inspected. The XML table structure was checked against these layouts. The source code is available as user-supplied text: the historical blanket claim that it

is unavailable is superseded. Its provenance is not independently authenticated publisher bytes. The intake audit is retained as historical evidence and its warnings were checked against the actual text; for example, `civ_bef` occurs at R243 rather than its cited R241. No historical manual or earlier draft supplied scientific design choices.

Live queries, dates, URLs and access limits are in `literature.md`. PMC/PubMed confirmed the article identity; the supplied full texts were readable. OpenAlex's exact-DOI request encountered a 429 daily-budget limit. BMJ main/related access returned 403, so a publisher correction or authenticated publisher code snapshot was not established. This is not evidence that no correction exists. Official Danish schema pages confirm field labels only. The CRAN `survival3.7-0` source archive and selected source/manual files were read without execution; its hash and access record are retained in the literature ledger.

Scientific fidelity and scope

Five final Atoms preserve the original analysis modules: population mortality, sibling comparison, surgery contrasts, treatment heterogeneity, and cause-specific mortality. The source-based reasons for these boundaries are in `source-map.md`. A single Molecule contains every Atom. Related coefficients, panels, outcomes and model variants remain together; routine linkage and coverage assertions are inside consuming protocols. There was no Atom-count target, candidate hand, ideation stage, novelty competition or review debate.

The original results are inventoried in their printed order. The machine catalogue contains 102 target groups, including all 32 Table 1 characteristic rows, the 14 published model blocks across Tables 2–3, all 29 supplementary product rows, figure/diagnostic and descriptive targets, and clearly identified source-code-only targets. Context and not-rebuilt material remain explicitly listed in `target-inventory.md`. The supplementary diagnostic remains one 14-panel figure within population mortality. The inventory distinguishes the paper's 703 younger-BSO deaths from a cohort population count.

The protocol preserves the source cohort, exclusions, prescription recipe, episode construction, time-varying exposure/covariates, age and calendar windows, Cox specifications, maternal-sibling comparison, surgery contrasts, cause-specific outcomes and source sensitivities. Shared construction is ordered pseudocode with exact code lists and logical upstream inputs. Formula aliases are explicitly defined. Every consuming output checks its source and input prerequisites before estimation. No new control, estimator, interaction, subgroup, sensitivity battery, scientific threshold, power calculation or repaired model was introduced.

D01–D20 in `deviations.md` are component-specific boundaries. The check covered commented prescription/clinical construction, all precomputed author files, `gravdb_steen` provenance, numeric date and age arithmetic, migration handling, dose units and future episode booking, progestin overlap, demographic rolling joins and imputation, adjustment-set discrepancies, route/regimen ties, sibling strata, retrospective older-surgery classification, cause horizon and record selection, smoking object/merge mismatches, `adj22`'s `m92` rather than `m222` reference, missing-to-one display replacement, and the dangling `rsvg` expression. Table 3's adjusted other-cause

values repeat the older-BSO adjusted values in the printed source; that correspondence was verified but is not a source-supported correction. Both an m92 substitution and a speculative m222 repair remain unauthorized.

Source-supported software clarification was narrow: pinned survival code verifies Efron ties, the covariance behavior for these calls, 95% summary intervals and the default cox.zph transform. The tmerge manual establishes pre-entry update and boundary-time behavior. Dropping unused baseline columns therefore does not, by itself, establish loss of pre-entry disease indicators. This clarification does not resolve missing incident histories, duplicate ordering, reported-model attribution or the unspecified ggcoxzph version. Exact plot-ready coordinates, curves, intervals and original axis ticks are required rather than a newly selected smoother.

Independent source-supported components remain registered when a neighboring component is blocked. Such conditional availability cannot repair a shared cohort gap or establish the missing main result. A future source clarification must identify its evidence and exact affected targets before execution. Merely naming an ambiguity, scientific extension or plausible fix does not authorize it.

Native authority and empirical separation

All five Atom registrations have atom_result=prereg, no observed evidence paths, no current local state and unfilled strength/shape. Each activation is TRUE. Shared input dependencies live inside protocols; planned bonds and Molecule interpretation rules describe evidence relationships. The native graph is derived with the existing repository renderer from these registrations, without new graph syntax or outcome-selected scheduling.

Boundary, partial and complete describe output coverage only. Operational failure remains no_result; it is neither a statistical null nor a completed evidence-backed boundary. Actual-result interfaces preserve component-specific estimates, intervals, denominators, mixed patterns and reasons for unavailability. No success, equivalence or significance-based replication scoring rule was added. Numerical compatibility ratings are not independent judgments of evidence or acceptance. All claims remain observational.

The 313 published reference cells are isolated in published-reference-values.csv and are never inputs to empirical result slots or plotting templates. All 292 manuscript result placeholders remain unfilled and bind to future governed outputs. Future results/atom-NNNNN paths are interfaces in the execution workspace outside this World; no such empirical outputs are present. The prepared manuscript includes the full source-aligned results tables, methods, conditional discussion, supplementary product and diagnostic presentation, and clearly marked source-code-only appendix.

Checks actually performed

Native package validation with the supplied allowed-table set and require_paper=True passed. Native read_bundle(..., birth=True) passed. Structural checks confirmed the exact slug mikkelsen-2026-hormone-therapy-native, complete

five-Atom Molecule membership, canonical activations, unique target IDs, every inventoried target, and 292 unique manuscript placeholders with matching Atom/output bindings. The source table parser was corrected during authoring after detecting an empty heading cell that shifted Table2 block labels; the final block row counts and displayed labels were checked against the PDF.

Both implemented figure templates rendered complete, partial and wholly missing synthetic input sets successfully. Synthetic fixtures and all six test PDFs remain outside the World in conversion/source-inspection/figure-tests. Additional interface checks rejected a numerical count attached to an unavailable flow node, a malformed residual interval, and an available panel lacking original axis ticks. The renderers do not fit models, infer absent counts or copy published values. Partial-input renders were visually inspected. An initial overlap between the residual figure's title/explanation text and panels was corrected in the renderer and the synthetic checks were rerun successfully. These are plotting/interface checks, not empirical tests or additional scientific diagnostics.

Source hashes and unchanged provenance/intake copies were checked. The final World contains only atoms/, molecules/ and prereg/; it excludes source PDFs, uploaded R, page screenshots, extraction caches, synthetic fixtures, generated figure PDFs, bytecode and archives. Validation scripts, logs, downloaded inspection material and synthetic renders remain outside the World. No source program, register analysis, full LaTeX/PDF build, sealing, ZIP creation, commit, push, release or external publication was performed.

Transcriber handoff

Preserve the five final Atom IDs, complete membership, all protocols and source gates, shared definitions, source audit, manuscript, result bindings and implemented figure sources. Native reader views are generated from those authorities and must not be treated as a competing specification. Transcriber and host own final formatting, LaTeX/PDF compilation and packaging. They may not resolve scientific conflicts or replace unavailable result slots with publication values. Successful formatting or package completion does not establish scientific executability of blocked targets or successful replication.

Source gaps and conflicts: no authorization to repair

P/S/R and line/page locators are defined in source-map.md. "Stop" means absent estimate and a named boundary for the affected target. It does not mean null effect. The full source code is available locally; the old blanket analysis-code-unavailable label is retired. Source-explicit rules and meaning-preserving engineering are the only implementation authority. This ledger does not authorize either member of a scientific conflict, parallel speculative variants, or new analyses.

ID	Checked finding and exact missing decision	Affected outputs and disposition
D01	P2-3 says 45th birthday/before45; R39-40,73-77 uses365.2-day years, rounds3 decimals, immigration<=25*365.2, exclusions<=start_age (R281-297). R75 has unquoted <code>as.IDate(1995-01-01)</code> , a numeric arithmetic expression, not a date string.	All cohort-derived targets: exact person inclusion/time-zero, boundary-day exclusions and age mapping unresolved. STOP_UNRESOLVED_SOURCE. Do not silently quote the date or replace birth-days/year length; code cannot run unchanged. Preserve paper's substantive target and code arithmetic in extraction.
D02	P3 censoring after emigration exceeding six months versus R51-53's first ordered row, U→I pairs >=180days, censor departure+180, plus status-derived immediate edato R72/76. No ordering guarantee in VNDS read; no return date for open departure.	Cohort, follow-up and all risk sets: need status-code dictionary, complete migration ordering and actual censor rule. Stop affected targets; do not pick 180days, six calendar months or departure itself as an invented common rule.
D03	R80-138 is commented, active R140/141 reads ht4/ht_pp; LMDB coverage starts1995 while P2 says since1994. R APK*VOLUME is called ddd; schema labels VOLUME only volume and NAME upper-case whereas code filters mixed-case pname.	All MHT targets: need author export product/volume metadata and lookback history; no generic ATC-DDD replacement, case-insensitive rewrite, or pre1995 no-use assumption.

ID	Checked finding and exact missing decision	Affected outputs and disposition
D04	R112–132 merges adjacent prescriptions into an episode; R384–389 assigns its total dose at its start. This may use future prescriptions for P’s prospective cumulative exposure. Separate-progestin match only catches endpoints inside estrogen interval; no covering-interval or competing-match decision.	Duration, exposure-type and cessation-dependent targets (potentially ever-use with affected eligibility/episodes): stop unresolved components. Do not distribute dose, fix overlap predicate, stockpile, or select a progestin arbitrarily. Preserve binary use only where fully determined independently of affected construction and shared cohort gates.
D05	op_all77_23 has no construction; unilateral cumsum R178 has no explicit chronological sort; LPR1 construction commented with setkey(lpr1) before assignment R184. P speaks discharge codes, R187 uses admission date. R222 counts diagnosis rows, P says contacts. Historical cancer export R212 and search-code prefix mapping absent.	Incident history, exclusions, surgery groups, health covariates, Table1: require original event/row and code mapping. No LPR-only cancer substitute, admission-for-operation-date proxy, retrospective sorting that changes scientific meaning or unique-contact repair.

ID	Checked finding and exact missing decision	Affected outputs and disposition
D06	R225 explicitly uses gravdb_steen because updated MFR was unavailable. Recipe producing gruppe1/8 , pregnancy rows and slutdato is absent; PARITET/BMI_MODER in FSV MFR are not equivalence proof.	Parity-adjusted models, parity descriptors and BMI/smoking sensitivity rows: STOP_UNAVAILABLE_INPUT until source-equivalent provenance. Other crude or count outputs do not require pregnancy covariates solely because R's large join includes them.
D07	BEF country alias fland, udda uddniv, civ_bef construction and ind1 year/population scope are not fully specified; R252 pooled inflation-adjusted quantiles need record scope and tie handling. R357/376 keys have extra columns and R359/378 roll against shorter keys. P4 mode imputation versus R445–448 hard-coded categories.	All affected demographic descriptors/adjusted fits: no origin proxy, new education crosswalk, widow mapping, rank/tie rule, annual quartiles, complete-case policy or guessed imputation population. Exact source mapping or invariant engineering proof required before output.

ID	Checked finding and exact missing decision	Affected outputs and disposition
D08	P1 abstract and T2 footnote list adjustment excluding civil status; P3 Covariates mentions civil status. R461 and all adjusted T2 fits include civst_mode; S figure1 visibly includes civst_mode. S figure verifies m02-like model identity, but does not explicitly correct the table footnote. P3 cardiovascular text omits four mediators, while T3 footnote/R756 omit valvular disease too; code also includes civil.	Published adjusted T2/T3 target attribution unresolved. Stop these adjusted outputs until exact reported model is reconciled. Preserve source-supported crude fits/summary siblings. Do not estimate both alternative adjustment sets. S1 PH is attributable to R m02 by visible term order but inherits unresolved model attribution for reproducing the published primary diagnostic.
D09	R302–325 splits at2023-01-01 but R342–344 reconstructs year_start from decimal_date(entry)+365.2 based elapsed years. R740/749/763/777 tests interval start rather than explicit stop truncation at2022-12-31. Code-only calendar filters R633/647 use2003, with no2003 split.	Cause-specific horizon and code-only calendar sensitivity risk sets: exact boundary handling unresolved. Do not silently add2003 splits or truncate differently. Main calendar covariates also require reconciliation if arithmetic crosses specified categories.
D10	R397–405 and418–426 assign strict cumulative maxima then nafill without by=pnr. Tie carry can cross persons, and first tied category is unspecified.	Route/regimen strata: stop targets requiring unresolved tie assignment; no alphabetical tie rule, cross-person inheritance or newly repaired by-person carry. First-use age targets retain separate status within Atom4.

ID	Checked finding and exact missing decision	Affected outputs and disposition
D11	P3 describes stratification by mother for sibling comparisons; adjusted R518 includes strata(family_id), crude R515 does not. R505 requires mother but not shared father; missing parent cleaning R45–47 depends on the full upstream person file.	Sibling adjusted model also D08; printed crude sibling HR's within-family interpretation unresolved. Counts/person-years can remain independent if cohort/family links resolved. Do not add strata to m31 or require full siblings.
D12	P3 surgical subgroups concern women who had undergone surgery. R583 overwrites bso=2 in all intervals for women whose eventual BSO age $\geq$ 55, including preoperative intervals. R670 does the same for hysterectomy $\geq$ 55; under55 hyst label includes earlier surgery too.	All $\geq$ 55 BSO summary/crude/adjusted targets and code-only older hyst subgroup have unresolved time origin; no retrospective exposure grouping or speculative postoperative repair. Hyst-any and BSO45–54 remain separately supported subject to shared gates; do not use them to claim older subgroup success.
D13	R714 constructs smoking-complete cpr8ryg and R715 sums it, but R717/720 fit cpr8bmi; R725 merges rt161 from BMI, not rt171.	Code-only smoking analysis: crude/adjusted/composite table target unresolved. Source-defined smoking-complete counts R715 may be retained alone if gravdb resolved. No changing cpr8bmi $\rightarrow$ cpr8ryg or rt161 $\rightarrow$ rt171. This is a checked source conflict, not an implemented corrected analysis.

ID	Checked finding and exact missing decision	Affected outputs and disposition
D14	R784 fits m222 for other mortality; R785 extracts m92 (older-BSO all-cause model). P T3 other-cause adjusted values exactly repeat T2 older-BSO adjusted values:0.88(0.67–1.15),0.72(0.48–1.07), verified against PDF pages6–7.	T3 adjusted other-cause targets unresolved/no estimate. Literal copying m92 is not an other-cause estimate; replacing it with m222 is not authorized. Cause-specific crude/summary siblings remain separately registered, conditional on their own mapping/horizon. No inferred correction from coincident printed values.
D15	R745 combines two cause files, orders person/underlying-cause string, takes first; cancer/cardiovascular classification overwrite precedence and unmatched→other. R746 uses event_pop for missing cause→other; R777 instead joins dar2. S says other causes, not unspecified causes. Source death versus status date fields differ.	Cause classification and missing/duplicate-record treatment unresolved when they determine endpoint. Need original version selection and missing-cause decision, not inferred all-cause subtraction or alternative classification. Known-cause siblings only if their risk/event mapping is independently source-determined.

ID	Checked finding and exact missing decision	Affected outputs and disposition
D16	R729/789 sets any missing adjusted estimate, adjusted text and crude text to 1. R463/etc binary coefficient assembly uses cbind/rbind with positional recycling. Several summary tables order on dropped hyst (R751/765/779), bso (R681/694).	Plot/table engineering must never turn failed fits into 1. Valid source reference categories may be explicitly labelled reference; named coefficient extraction is meaning-preserving only when it returns the already specified coefficient of the already authorized model. Nonreference absent estimate stays absent. Removing a dropped-column sort is layout-only when scientific row identity is unchanged; it does not fix a subgroup/model.
D17	R947 trailing rsvg:: is incomplete/dangling and may consume following tokens unexpectedly; R946 calls rsvg for a jpg file after prior PDF exports.	Whole script cannot be represented as runnable. Reimplementing flowchart geometry from P fig1 and validated counts is engineering; no empirical R execution needed. Keep rsvg problem distinct from unavailable scientific counts.

ID	Checked finding and exact missing decision	Affected outputs and disposition
D18	S figure1 has14 labelled covariate panels beginning ht, includes civst_mode; R465–466 plots m02. R480 computes cox.zph(m12) but does not plot it. P3 accepts PH visually; S displays small/rounded-to-zero test P values. No formal acceptance algorithm or fitted-smoother software version is supplied for ggcoxzph.	Rebuild S1 only from exact m02/cox.zph/ggcoxzph equivalent plot-ready output once model/software provenance is resolved. Do not substitute duration model, infer PH violation from rounded P alone, create time interactions, or add a threshold. Retain plots/P values and qualitative assessment without automatic “passed”.
D19	S2 says proportion of MHT used; R430 instead sums risktime on rows carrying nonmissing episode ATC (not all MHT DDD).	S2 each product/form/regimen proportion and associated most-common-product prose unresolved denominator. No DDD-weighted or person-weighted replacement. Product classification definitions remain extracted.
D20	R699/705 explicitly supports BMI<30 versus>=30 complete-observed-interval sensitivity; calendar, surgery-adjusted and age-hyst sensitivities are in R632–725 but absent from P/S results. R832 density and R835–840 commented single-prescription/statin explorations have no published result targets.	Retain code-only model specifications with no published numerical target, inside original modules; display them separately from published Table2. Density and commented exploration are explicitly not rebuilt/context, not additional experiments. BMI30 is source-explicit, not a converter-added threshold.

Resolved engineering and source-supported clarifications

- Intake note’s `civ_bef` locator241 is historical; checked active read is R243. All969 lines were read, CRLF bytes preserved in source. No code is redistributed in the World.
- S1 panels identify binary `m02` rather than duration `m12`: visual `Beta(t)` for `ht`, covariate order and R465–466 agree. This resolves which diagnostic is being inventoried; it does not reconcile D08.
- Prescription primary ever-use persists after cessation (P2–3 and R427 agree). Contraindication censoring is a sensitivity, not the main analysis (P3/R486 agree).
- Label spelling, filenames, long-form output columns, explicit reference labels and numbered missing-display explanations are engineering. No numerical scientific quantity is filled or recomputed in plotting code.
- The abstract’s “703 women who underwent BSO” wording is clarified by P5 and R597: the count refers to women in that group who subsequently died. It is a death-count target, not the full surgery cohort size.

No new scientific choice was authorized during conversion. Outstanding conflicts remain visible through transcription and compilation. A successful build cannot resolve them.

New source intake: Mikkelsen supplement2.txt

Received directly from the user on 2026-09-05 after the Slack files-pri URL required authentication. The uploaded attachment was copied byte-for-byte; no R code was executed and no referenced data path was opened.

- Attachment: `/tmp/codex-remote-attachments/01a05b97-1ecb-72d0-84c7-4ed89f71b9d8/AB1DD9EE-5C35-42E4-AB74-A12E547A281B/1-supplement2.txt`
- Retained local source: `supplements/mikkelsen-supplement2.txt`
- SHA256: `a386ae0480f42705d5d8cf0b850f703123e962265e4d46715cecb2ecac0cfeea`
- Size: 65,619 bytes; 969 newline-terminated lines, CRLF retained.
- File identifies R 4.4.1 and a last change-log entry of 2025-08-16, v2.1.
- Provenance: user-supplied study code; independent publisher-file provenance is not yet verified. Its MHT definitions, 1950–1977 female cohort, age-45 entry and 2023-07-31 end match the target study’s setting.

What this resolves

The previous statement that no analysis code was available is superseded for this local conversion. The earlier Europe PMC supplement snapshot still does not itself contain the program; the code was obtained separately from the user. Do not carry forward a blanket `analysis_code_unavailable` limitation.

The file specifies raw/derived source construction, prescription episodes and cumulative volume, interval splitting, time-varying covariates, main Cox models, sibling stratification, contraindication/cessation censoring, formulation and regimen, surgery/age/calendar subgroups, BMI/smoking checks, cause-specific mortality and

descriptive tables/flowchart generation. Read all 969 lines and map the relevant sections to the exact paper/table/Atom targets.

Important remaining questions (source inspection, not empirical findings)

1. Large prescription and data-building sections are commented out, while the active program reads precomputed files, e.g. `ht4.csv` (line 140), `ht_pp.csv` (141), `op_all77_23.csv` (172), `civ_bef.csv` (241) and others. This is not a self-contained reproduction environment. Inventory each missing upstream recipe/file; distinguish commented source recipes from executable sections.
2. The pregnancy source is `gravdb_steen.csv` (line 226), with a comment that the updated Medical Birth Register was unavailable. Map the construct to current FSV carefully: access to MFR does not establish equality to this derived source. Parity/BMI/smoking provenance and timing need explicit reconciliation.
3. The smoking block constructs `cpr8ryg` (714), but `m171/m172` use `cpr8bmi` (717/720), and `rt172` merges `rt161` (725). These are visible object-reference discrepancies, not proof of which implementation produced published results.
4. The other-cause mortality block fits `m222` (784) but `adj22` extracts `summary(m92)` (785). Register an explicit discrepancy, distinguishing literal code replay from a documented method-consistent correction; never silently change it.
5. A trailing standalone `rsvg::` expression near the flowchart block is incomplete. Do not claim the whole uploaded script can run unchanged end-to-end.
6. The file contains source-specific ordering, time rounding (365.2-day years, precision 3), product volume, ties/rolling joins and handling of absent estimates. Audit these rather than replacing them with convenient defaults. In particular, filling missing adjusted estimates with 1 in display assembly must never turn a failed replication fit into a null finding.

The new converter must retain uncertainties and choose explicit, pre-execution dispositions. Do not fabricate upstream files or claim source-code agreement establishes scientific validity. Preserve this intake note in `docs/replication/`.

Source-to-Atom audit

Source hierarchy and identity

P = Mikkelsen AP, Bergholt T, Lidegaard Ø, Scheller NM. *Menopausal hormone therapy and long term mortality: nationwide, register based cohort study*. BMJ2026;392:e085998, doi:10.1136/bmj-2025-085998. P locators are printed PDF pages1–10. S = supplied `prea085998.ww.pdf`, four pages: contents p1, definitions p2, product-use table p3, residual figure p4. R = user-supplied `mikkelsen-supplement2.txt`, 969 newline-terminated CRLF lines; R line numbers count blank/commented lines. All seven supplied source hashes were checked against `source-provenance.yaml`, which is copied unchanged. Source PDFs, code and full-text caches are outside this World.

The complete paper, all supplement pages and all969 R lines were read. The local article Markdown is a malformed table extraction and not the authority for column alignment. The JATS XML identifies PMC12915068/PMID41708152/DOI and carries only the supplementary PDF link, repeated in associated-data material; it does not authenticate the separately supplied code. Table identities and public reference cells were extracted from its tbl1/tbl2/tbl3 structure and checked against PDF layout. Critical Table2/Table3 and S figure1 pages were visually inspected. Published numerical cells are isolated in published-reference-values.csv and never bound to current results. The study's P2 visual abstract duplicates original targets; it is retained in inventory as context, not recreated as branded artwork.

Historical README and resource-classification warnings about analysis-code absence describe an earlier accessible bundle. They are superseded for this run by the user-supplied R code. The intake audit is copied unchanged as mikkelsen-supplement2-intake.md; its warnings are leads, not scientific authority. The corrected line location for civ_bef is R243, not241. No historical deck, earlier draft or replication manual was adopted as a design. Independent publisher provenance for R is unresolved: BMJ live main/related pages returned403; the code's internal R4.4.1 and 2025-08-16 v2.1 labels are self-description, not authentication.

Final scientific units

	Original coherent analysis and covered targets	Why this boundary is scientifically meaningful
Final Atom atom-000001	Population binary/duration all-cause association; cohort flow, uptake, age55 characteristics; contraindication and cessation checks; S2 use distribution; S1 PH; code-only calendar, surgery-adjusted, BMI and smoking checks	These assess or describe the same nationwide association. Routine joins, diagnostic plots and each duration coefficient are not independent questions. Count/PY/crude/adjusted and check-specific availability remain distinct inside this Atom.

Final Atom	Original coherent analysis and covered targets	Why this boundary is scientifically meaningful
atom-000002	P/Table2 sibling comparison and its population size	The source deliberately changes the comparison to discordant maternal families to address shared familial confounding. This has independent interpretation and selection limits. Crude/adjusted models stay together, with their strata discrepancy explicit.
atom-000003	P/Table2 hysterectomy and BSO age contrasts, BSO death-age descriptions; code hysterectomy-age variants	The source asks about mortality in clinically distinct surgery populations. Related surgery and age contrasts stay together, including overlapping memberships. Their source timing conflicts and partial availability are preserved without overmerging the nationwide association.
atom-000004	P/Table2 dominant route, progestogen regimen and initiation-age strata	These jointly describe heterogeneity in exposure, with separate source comparisons and no invented head-to-head model. Tie/episode gaps need not erase independently determinable initiation-age output.

Final Atom	Original coherent analysis and covered targets	Why this boundary is scientifically meaningful
atom-000005	P/Table3 cardiovascular, cancer and other mortality, cause proportions	The source decomposes mortality using a common cause-register horizon. Related outcomes stay together, with independent estimate/CI/status records. Other-cause adjusted corruption does not become a separate experiment and cannot erase its faithful crude or other-cause siblings.

No numerical Atom target was imposed. Each distinct question has an actual-result interface and downstream interpretation. No figure, panel, coefficient, data linkage or software operation automatically became an Atom. The five-member Molecule owns the whole roster; Atom registrations own protocols and coverage regions. Coverage states are not scientific evidence classifications.

Source rules, translation and limits

The complete consequential instruction map is the numbered shared construction in data-contract.md, exact clinical-codes.yaml, each Atom’s protocol and D01–D20 in deviations.md. **target-catalog.yaml** gives every source row/group and output identity; target-inventory.md presents them in original result order. **schema-map.csv** supplies92 exact candidate fields, source documentation paths, official URLs and coverage; none is marked mounted/readable/validated. A candidate mapping is not a licensed substitute for a missing original register construction.

Rule family	Source-explicit evidence	Engineering translation allowed	Unresolved scientific boundary
Eligibility/time	P2–3, S1; R39–77,279–297	Typed dates/person keys and source-exact joins once established	D01/D02 birthdays, entry inclusivity, numeric date expression, migration/status history

Rule family	Source-explicit evidence	Engineering translation allowed	Unresolved scientific boundary
Prescription supply/episodes	P2–3, S2; commented R80–138, active reads140–141	Name output columns; preserve source grouping/splits	D03/D04 export dose units, historical coverage, grouping order, future dose booking, separate progesterin overlap
Clinical history	S1; R174–222	Lossless verified key/date/code translation	D05 raw/derived cancer and surgery files, laterality/order, contact vs diagnosis unit
Covariates	P3–4; R224–380,444–448	Exact crosswalk only if documented; no new category or sample	D06 gravidb equality; D07 birth-country/education/civil/income mappings, mode and ties
Cox	P3–4/T2/T3; R458–784	Named extraction of an authorized coefficient and source95%CI, fixed output paths	D08 printed adjustment attribution; D11 strata; D12 risk time; D13/D14 wrong model objects
Mortality causes	P3, S1; R740–799	Exact case/alias mapping after version/date verification	D09 cutoff; D15 primary cause version and missing classification
Figures/tables	P fig1, S fig1; R465–466,729/789,842–960	Python render of approved counts/residuals/cumulative layout, explicit missing reasons, references labelled	D16 no missing→1; D17 dangling rsvg; D18 source plot implementation; D19 product-use denominator

The native graph is canonical. All activations are TRUE: the units share logical input construction but no unit needs a prior unit’s realized statistical result to run. Bonds describe planned interpretation relationships, and the existing graph’s dotted

interpretation edges compose those findings in the Molecule. Shared tables are input dependencies inside each protocol. A **done** interpretation condition includes a completed evidenced boundary and therefore never promises that an estimate exists. No new graph syntax, outcome-selected schedule, significance routing or extra synthesis Atom is introduced.

Source scope outside the rebuilt result set

P’s external trials/cohorts, symptom/usage background, clinical-guideline discussion, original funding/ethics and prior literature are context, not fresh replication experiments. Their factual claims are attributed to the original paper and not asserted as current clinical guidance. R’s active unreported calendar/BMI/surgery adjustments are retained as source-code-only specifications and no published numbers are invented for them. Source-conflicting smoking/hyst variants remain unresolved targets. R’s unused nonfatal diagnosis objects, duration-density inspection and commented one-prescription/statin exploration are explicitly not rebuilt. They do not authorize additional science.

Every target remains in the inventory even where an input is unavailable or paper and code disagree. Transcriber must preserve these gaps and the final IDs. This is conversion with a self-check, not independent scientific review or original-author approval.

Original target inventory

The order below follows the article Results narrative and the original within-table block/row order, then discussion-only quantities, supplementary definitions/product table/diagnostic, followed by code-only targets in R order. Abstract and visual-abstract repetitions are aliases of these same targets. Published reference values are isolated in published-reference-values.csv; they are not current results. Every row/category has its own eventual status even when an inventory group contains several cells. No empirical result is filled here.

Target	Original locator and full scope	Final Atom	Rebuild disposition
<code>cohort_flow</code>	P Figure1 p4; R279-299,806- 813: Eligible/exclusion- union and seven overlapping reasons; included/end-use groups/duration counts	atom-000001	Registered source target; current values unfilled; consuming protocol and D01- D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
followup	P Results p4; R817: Follow-up median/IQR in years	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
uptake	P Results p4; R819–827: End-use N/percent;1950,1960,1970 birth-year uptake; exposed-duration median/IQR; final duration proportions	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
cohort_deaths	P abstract; R822–823: Death N/percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
landmark_uptake	P Results p4; R957–960: MHT at55 in2004–06 and2021–23	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
landmark_n	P Table1 header; R950–955: At55 N by past/present versus never use	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
t1_03	P Table1 p5: Denmark; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_04	P Table1 p5: Other; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_05	P Table1 p5: Unknown; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_06	P Table1 p5: Mean parity (SD); rows: Past/present at55, No use at55; components: mean_sd	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_08	P Table1 p5: 0; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_09	P Table1 p5: 1-2; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
t1_10	P Table1 p5: 3; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_12	P Table1 p5: Unmarried; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_13	P Table1 p5: Married; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_14	P Table1 p5: Divorced; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_15	P Table1 p5: Unknown; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_17	P Table1 p5: Primary or lower secondary school; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
t1_18	P Table1 p5: Upper secondary school; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_19	P Table1 p5: Bachelor’s degree or equivalent; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_20	P Table1 p5: Master’s degree or equivalent; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_21	P Table1 p5: Unknown; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_23	P Table1 p5: 1st; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_24	P Table1 p5: 2nd; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
t1_25	P Table1 p5: 3rd; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_26	P Table1 p5: 4th; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_28	P Table1 p5: 2002-2008; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_29	P Table1 p5: 2009-2015; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_30	P Table1 p5: 2016-2023; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_31	P Table1 p5: Diabetes; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
t1_32	P Table1 p5: Hypertension; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_33	P Table1 p5: Hypercholes- terolemia; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_34	P Table1 p5: Atrial fibrillation; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_35	P Table1 p5: Heart failure; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_36	P Table1 p5: Valvular disease; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
t1_37	P Table1 p5: Three or more hospital contacts between age 44 and 45 years; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_38	P Table1 p5: Earlier hysterectomy; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t1_39	P Table1 p5: Earlier bilateral oophorectomy; rows: Past/present at55, No use at55; components: n_percent	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t2_primary	P Table2 p6; R456–469: Primary analysis; rows: Never used, Past or present use; components: summary, crude, adjusted	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t2_duration	P Table2 p6; R472–483: Primary analysis stratified by duration of use; rows: 0 years, <1 year, 1-2.9 years, 3-4.9 years, 5-9.9 years, 10 years; components: summary, crude, adjusted	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
t2_contraindications	Bn Table2 p6; R486–500: Sensitivity analysis: censored at MHT contraindication†; rows: Never used, Past or present use; components: summary, crude, adjusted	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
sibling_population	B Results p4; R511–512,829– 830: Sister N and share of original cohort	atom-000002	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t2_sibling	P Table2 p6; R503–523: Sensitivity analysis: sibling analysis; rows: Never used, Past or present use; components: summary, crude, adjusted	atom-000002	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t2_cessation	P Table2 p6; R526–542: Sensitivity analysis: ‘as treated’ analysis (censored at MHT cessation); rows: Never used, Past or present use; components: summary, crude, adjusted	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
t2_hyst	P Table2 p6; R569–580: Subgroup of women who had undergone hysterectomy§; rows: 0 years, <5 years, 5 years; components: summary, crude, adjusted	atom-000003	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t2_bso_45_54	P Table2 p6; R583–599: Subgroup of women who had undergone bilateral oophorectomy between age 45-54 years§; rows: 0 years, <5 years, 5 years; components: summary, crude, adjusted	atom-000003	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t2_bso_ge55	P Table2 p6; R603–613: Subgroup of women who had undergone bilateral oophorectomy at age 55 years or older§; rows: 0 years, <5 years, 5 years; components: summary, crude, adjusted	atom-000003	Risk-set target blocked D12; no speculative postoperative repair.

Target	Original locator and full scope	Final Atom	Rebuild disposition
bso_deaths	P p5; R597: Death count among BSO45–54 risk set, not BSO population N	atom-000003	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
bso_death_age	P p5; R598–599: Median/IQR tstop among deceased women by use at death	atom-000003	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t2_route	P Table2 p6; R545–554: Stratified analysis: by treatment form most used; rows: Never used, Oral, Transdermal, Other formulation; components: summary, crude, adjusted	atom-000004	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t2_regimen	P Table2 p6; R557–566: Stratified analysis: by progestogen regimen most used; rows: Never used, Oestrogen monotherapy, Oestrogen+cyclic progestogen, Oestro- gen+continuous progestogen; components: summary, crude, adjusted	atom-000004	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
t2_initiation	P Table2 p6; R616–630: Stratified analysis: by age of initiation; rows: Never use, 45-51 years, 52-56 years, 57 years; components: summary, crude, adjusted	atom-000004	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
cause_proportions	P p5; R797–799: Cardiovascu- lar/cancer/other fractions of classified death counts	atom-000005	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t3_cardiovascular	P Table3 p7; R740–799: Cardiovascular mortality†; rows: 0 years, <5 years, 5 years; components: summary, crude, adjusted	atom-000005	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t3_cancer	P Table3 p7; R740–799: Cancer mortality‡; rows: 0 years, <5 years, 5 years; components: summary, crude, adjusted	atom-000005	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
t3_other	P Table3 p7; R740–799: Other mortality‡; rows: 0 years, <5 years, 5 years; components: summary, crude, adjusted	atom-000005	Adjusted output blocked D14 (m92 vs m222); crude/count siblings retained subject to D09/D15 and shared gates.

Target	Original locator and full scope	Final Atom	Rebuild disposition
contraindication	Deaths Discussion p8; R489: Percent of deaths preceded by contraindication	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
end_age	P limitations p8; R final tstop: Median end-of-follow-up age	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
missing_demographics	Methods p4; R963–967: Final-row educa- tion/civil/country/income missing percentages	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.
s1_definitions	S table1 p2: Definition source incorporated into shared protocol; not an empirical experiment	atom-000001	Context / explicitly not rebuilt as a new empirical analysis; retained in source audit.
s2_01	S Table2 p3 row1; R429–436: G03CA03 / Oestrogen / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_02	S Table2 p3 row2; R429–436: G03CA03 / Oestrogen / Transdermal; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.

Target	Original locator and full scope	Final Atom	Rebuild disposition
s2_03	S Table2 p3 row3; R429–436: G03CA03 / Oestrogen / Injection; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_04	S Table2 p3 row4; R429–436: G03CA03 / Oestrogen / Nose spray; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_05	S Table2 p3 row5; R429–436: G03CA03 / Oestrogen + Separate cyclic progestin / Transdermal; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_06	S Table2 p3 row6; R429–436: G03CA03 / Oestrogen + Separate cyclic progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_07	S Table2 p3 row7; R429–436: G03CA03 / Oestrogen + Separate cyclic progestin / Injection; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.

Target	Original locator and full scope	Final Atom	Rebuild disposition
s2_08	S Table2 p3 row8; R429-436: G03CA03 / Oestrogen + Separate cyclic progestin / Nose spray; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_09	S Table2 p3 row9; R429-436: G03CA03 / Oestrogen + Separate continuous progestin / Transdermal; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_10	S Table2 p3 row10; R429-436: G03CA03 / Oestrogen + Separate continuous progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_11	S Table2 p3 row11; R429-436: G03CA03 / Oestrogen + Separate continuous progestin / Injection; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_12	S Table2 p3 row12; R429-436: G03CA03 / Oestrogen + Separate continuous progestin / Nose spray; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.

Target	Original locator and full scope	Final Atom	Rebuild disposition
s2_13	S Table2 p3 row13; R429–436: G03CA04 / Oestrogen / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_14	S Table2 p3 row14; R429–436: G03CA04 / Oestrogen + Separate cyclic progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_15	S Table2 p3 row15; R429–436: G03CA04 / Oestrogen + Separate continuous progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_16	S Table2 p3 row16; R429–436: G03CA53 / Oestrogen / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_17	S Table2 p3 row17; R429–436: G03CA53 / Oestrogen + Separate cyclic progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_18	S Table2 p3 row18; R429–436: G03FA01 / Oestrogen and continuous progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.

Target	Original locator and full scope	Final Atom	Rebuild disposition
s2_19	S Table2 p3 row19; R429–436: G03FA01 / Oestrogen and continuous progestin / Transdermal; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_20	S Table2 p3 row20; R429–436: G03FA11 / Oestrogen and continuous progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_21	S Table2 p3 row21; R429–436: G03FA12 / Oestrogen and continuous progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_22	S Table2 p3 row22; R429–436: G03FA15 / Oestrogen and continuous progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_23	S Table2 p3 row23; R429–436: G03FA17 / Oestrogen and continuous progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_24	S Table2 p3 row24; R429–436: G03FB01 / Oestrogen and cyclic progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.

Target	Original locator and full scope	Final Atom	Rebuild disposition
s2_25	S Table2 p3 row25; R429–436: G03FB05 / Oestrogen and cyclic progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_26	S Table2 p3 row26; R429–436: G03FB05 / Oestrogen and cyclic progestin / Transdermal; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_27	S Table2 p3 row27; R429–436: G03FB06 / Oestrogen and cyclic progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_28	S Table2 p3 row28; R429–436: G03FB09 / Oestrogen and cyclic progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s2_29	S Table2 p3 row29; R429–436: G03FB11 / Oestrogen and cyclic progestin / Oral; percentage of use	atom-000001	Definition retained; proportion no- estimate pending denominator D19 and exposure inputs.
s1_schoenfeld	S figure1 p4; R465–466: 14 binary-primary- model term panels; R480 duration diagnostic unplotted	atom-000001	Registered source target; current values unfilled; consuming protocol and D01– D20 specify exact component stops.

Target	Original locator and full scope	Final Atom	Rebuild disposition
code_calendar_pre	R663-645: Code-only calendar interval- start<2003; rows: Never used, Past or present use; components: summary, crude, adjusted	atom-000001	Source-code-only specification; no published numerical comparator. Preserve explicit conflict/no- estimate components.
code_calendar_pos	R663-659: Code-only calendar interval- start>=2003; rows: Never used, Past or present use; components: summary, crude, adjusted	atom-000001	Source-code-only specification; no published numerical comparator. Preserve explicit conflict/no- estimate components.
code_surgery_adj	R662-667: Code- only additional hyst/BSO adjustment; rows: Never used, Past or present use; components: summary, crude, adjusted	atom-000001	Source-code-only specification; no published numerical comparator. Preserve explicit conflict/no- estimate components.
code_hyst_lt55	R669-682: Code-only hysterectomy- age<55; comment45-54 unresolved; rows: 0 years, <5 years, 5 years; components: summary, crude, adjusted	atom-000003	Source-code-only specification; no published numerical comparator. Preserve explicit conflict/no- estimate components.

Target	Original locator and full scope	Final Atom	Rebuild disposition
code_hyst_ge55	R684-695: Code-only hysterectomy- age>=55; rows: 0 years, <5 years, 5 years; components: summary, crude, adjusted	atom-000003	Source-code-only specification; no published numerical comparator. Preserve explicit conflict/no- estimate components.
code_bmi	R698-710: Code-only BMI observed-interval sensitivity, threshold30; rows: Never used, Past or present use; components: summary, crude, adjusted	atom-000001	Source-code-only specification; no published numerical comparator. Preserve explicit conflict/no- estimate components.
code_smoking	R713-725: Smoking-complete summary; model/merge conflicts, no repaired model; rows: Never used, Past or present use; components: summary, crude, adjusted	atom-000001	Source-code-only specification; no published numerical comparator. Preserve explicit conflict/no- estimate components.
code_density	R832: Unpublished duration density inspection; not rebuilt	atom-000001	Context / explicitly not rebuilt as a new empirical analysis; retained in source audit.
code_commented_explo	R835-840: Commented single- prescription/statin exploration; not rebuilt	atom-000001	Context / explicitly not rebuilt as a new empirical analysis; retained in source audit.

Target	Original locator and full scope	Final Atom	Rebuild disposition
code_unused_outcome	P198,208– 210,218–219: Unused nonfatal/all- cancer intermediate objects; no outcome models added	atom-000001	Context / explicitly not rebuilt as a new empirical analysis; retained in source audit.
visual_abstract	P p2: Duplicate popula- tion/sibling/BSO results and descriptive cohort; not independently rebuilt as branded graphic	atom-000001	Context / explicitly not rebuilt as a new empirical analysis; retained in source audit.
external_comparison	P p1–2,7–8 and references: WHI, observational cohorts and policy background; no new experiments or current clinical recommendation	atom-000001	Context / explicitly not rebuilt as a new empirical analysis; retained in source audit.

Table1 includes every printed category, not merely its variable headings. Table2 includes every reference and nonreference row and summary/crude/adjusted field. Table3 includes all three causes and duration groups. S2 includes all printed ATC–regimen–form rows in printed order. Supplementary table1 is a definition source, not an unperformed experiment. Source labels such as ever-use and never-use refer to their table-specific time semantics; neither the landmark table nor final cohort flow changes Cox exposure timing.

The source P2 visual abstract repeats primary, sibling and BSO hazard-ratio/count targets, including the different units of incidence rates. Its protective/harmful annotations and source conclusions are not new preregistered decision rules. The current manuscript uses observational wording and no projected finding. P1 abstract death/uptake quantities map to cohort_deaths/uptake; its BSO703 wording maps to bso_deaths, as clarified in P5. P Methods missing percentages map to missing_demographics; P8 contraindication-death percentage and end age remain explicitly inventoried.

Read-only inventory of original program inputs

These are names and locators from supplied R source, not files opened by the converter. No source data file is asserted present or recreated. The author's E: work paths have no runtime meaning here. Source recipes are read-only evidence; commented code is not proof of a successful original build.

Input	R lines	Required content / recipe status	Consumers and disposition
t_person2022.csv	45	pnrm, pnrm, pnrf, c_kon, c_status, fdat, raw extraction/version absent	All children, fdat, d_status_hen_start; sibling targets; D01/D02 and person-history mapping required
vnds2022.csv	50	pnrm, haend_dato, indn, fdat, input ordering unstated	All folders; up; D02
bef.csv → bef2.csv	59–65	commented unique pnrm/ie_type/fland recipe; active read sorted by fland	Birth country and adjusted models; D07
hc2022.csv → ht4.csv	80–140	commented prescription extrac- tion/aggregation; ht4 contains pnrm, fdat, htstart, htstop, htage, htform, streng, httype, htname, ddd	All exposure- dependent targets; D03/D04. Do not execute uncommented design
ht_pp.csv	136–141	commented earliest 90-day- prolonged cessation from ht4	As-treated sensitivity; D04 and source sequence required
dm1.csv	144–147	commented earliest A10 except Wegovy Flextouch, diabetes_age	Diabetes adjustment and Table1; preserve source filter, verified product alias required
hyperten3.csv	150–160	commented class assignments then first second distinct class, ht_age	Hypertension adjust- ment/Table1; do not count overlapping ATC classes differently

Input	R lines	Required content / recipe status	Consumers and disposition
statin1.csv	163–168	commented C10AA/C10AB first prescription, statin_age	Hypercholesterolaemia adjustment/Table1
op_all77_23.csv	172	pnr,okode,odto; no construction for unified old/new register procedures	All baseline BSO exclusions and surgery analyses; D05
lpr77_23_ab.csv → lpr1.csv	183–190	commented join, admission-age, per- son/date/diagnosis sort; dangling preassignment setkey(lpr1)	Exclusions/clinical covariates/contact count; D05 requires exact merged hospital history
loop.csv	201–204	commented earliest C03C prescription; pnr,eksd	Combined heart failure definition; requires diagnosis and prescription
cancer.csv	212	pnr,search,d_diagnosis raw source extraction absent	Merge to breast/gynaecological and all-cancer history; historical Cancer Register cannot be silently omitted
gravdb_steen.csv	225–230	pnr,slutdato,gruppe, source pregnancy database construction missing	Parity; BMI/smoking; D06. Current MFR permission does not make this file equivalent
udda2022.csv	238–240	pnr,uddniv,uddniv1, uddniv crosswalk not supplied	Time-dependent education; D07
civ_bef.csv	243–245	pnr,civst,civ_start; missing historical merged-file recipe	Civil state/adjustment; D07. Intake note's241 locator corrected here


```

/home/shaoerzhuo/anaconda3/bin/python -B docs/execution/run.py --help
/home/shaoerzhuo/anaconda3/bin/python -B docs/execution/run.py \
  --atom atom-000001 --world "$WORLD" --inputs "$INPUT_MANIFEST" \
  --output "$ACTION/results/atom-000001" \
  --dispatch-policy activation_respecting
/home/shaoerzhuo/anaconda3/bin/python -B docs/execution/run_tests.py \
  --output "$ABSOLUTE_FRESH_SCRATCH_DIRECTORY"

```

WORLD, INPUT_MANIFEST, ACTION, and ABSOLUTE_FRESH_SCRATCH_DIRECTORY above are caller-supplied paths, not presumed mounts. `--output` must be an empty directory outside the World. `--inputs` accepts a JSON file or a directory with `manifest.json`. Exit 0 includes a completed source boundary; it does not mean an estimate exists. Exit 2 indicates rejection or an operational `no_result`. Tests keep Matplotlib's cache under their scratch output; the standalone renderer uses its output directory when no caller-supplied MPLCONFIGDIR exists.

Every execution (including a governed boundary assessment) now requires the external controller trust anchor described below. `--help` requires no anchor or data. `run_tests.py` creates its own explicitly synthetic controller registry in the requested scratch directory; this grants no real execution acceptance.

Tested: Python 3.12.4, NumPy 1.26.4, pandas 2.2.3, SciPy 1.15.2, PyYAML 6.0.2, Matplotlib 3.10.0, statsmodels 0.14.4, pypdf 6.11.0; R 4.1.2, survival 3.8.3, jsonlite 2.0.0. The R package `data.table` 1.18.2.1 is installed but the adapter does not use it. Frozen source versions are **R 4.4.1**, **survival 3.7-0**, **data.table 1.16.2**. The installed R engine is used only for labeled synthetic checks; `Engine(..., synthetic=False)` enforces the R/survival pin. Those checks are actual numerical execution, but do not validate the unavailable pinned engine. `survminer` and its source version are unavailable. No source-compatible Schoenfeld smoother is asserted. `pdftoppm/pdftotext` are available for local PDF inspection. Publication Helvetica Neue resources were not present in the searched repository resources; local figure tests use the fixed template's documented DejaVu Sans fallback. Supply the host `style_dir` containing its four Helvetica Neue files for publication rendering.

Runtime input contract

A minimal, runnable pre-data assessment manifest is:

```

{"purpose": "governed", "datasets": {}, "definitions": {}, "source_evidence": [], "predecessors": []}

```

`run_tests.py` creates complete synthetic manifests and input CSVs in its scratch directory. A dataset descriptor has this shape (the path is deliberately a placeholder):

```

{
  "purpose": "synthetic",
  "datasets": {
    "intervals": {
      "path": "/EXPLICIT_CALLER_PATH/intervals.csv",
      "sha256": "CONTENT_SHA256",
      "origin": "synthetic_fixture",
      "tables": [],

```

```

    "source_locator": "Synthetic lower-stage fixture for R382/R444-R448
    ↪ logical columns",
    "columns": {"pnr": "string", "tstart": "number", "tstop": "number", "event": "
    ↪ number", "ht": "number", "ht_sum": "number"}
  }
},
"definitions": {}, "source_evidence": [], "predecessors": [],
"rscript": "/EXPLICIT_CALLER_PATH/Rscript"
}

```

CSV columns must match the declared `columns` map exactly; types are `string` or `number`. Empty cells are missing; literal NA strings are not silently parsed as missing civil categories. Person and parent keys are strings, preserving leading zeros and large IDs. Synthetic person keys must begin `SYN-`, every synthetic dataset must declare `origin: synthetic_fixture`, and all generated artifacts are marked synthetic. Marking a real dataset synthetic would violate this input contract; the label is not a deidentification operation.

Every declared dataset requires a 64-hex SHA256, validated before dispatch or row access, including governed runs blocked at D01/D02. Hex is normalized to lowercase. Omit an absent optional dataset from `datasets`. A valid declaration does not prove that its path exists or that its content matches: `load_frames` checks actual bytes only when reading an authorized input. Governed boundaries read no such bytes, so their input identity describes declared inputs only.

Accepted dataset names are `intervals`, `cause_intervals`, `cohort`, `persons`, `causes`, `changes`, and `episodes`. These are **logical interfaces, not invented physical FSV tables**. Their exact fields and provenance are below. Missing an optional dataset or column stops only its consumers. Corrupt hashes or malformed manifests reject the invocation instead of producing a statistical null. Runtime table declarations must be within the unchanged 531-table package boundary. The code does not expand the scientific table scope or select new tables from it.

`definitions` and `source_evidence` are provenance, not executable rules. A source-evidence item requires `locator` and `sha256`; it records source-backed dictionaries or extraction documentation. It cannot clear D01–D20, select an adjustment formula, change classification/version, choose a missing-value rule, or supply an unknown date/parameter. Physical FSV-to-author-export equivalence, including PNR12/PNR, NAME/pname, VOLUME/DDD, and gravdb provenance, is still unestablished. The raw construction functions deliberately accept the exact author aliases after such equivalence, never guessed physical files. No new scientific registry schema is introduced.

Scientific `source_locator`, origin and dictionary metadata must be portable (for example R line ranges, a DOI/page/field, or an official schema URL). Absolute machine paths are accepted only in runtime path fields, and rejected inside scientific identity metadata; relocations cannot change the input identity.

`predecessors` contains objects `{"id": "HOST_ACCEPTED_RESULT_ID"}`. Paths, hashes and acceptance come from the externally anchored controller registry, never from an action manifest or the result's own `accepted` Boolean. Only host-accepted, nonsynthetic evidence with matching frozen material, logical input identity and an

approved implementation is eligible. Its aggregate hashes, result digest and controller-bound result byte hash must verify. The executor never marks its own output accepted. Native `registered_state` belongs to the host's actual scientific result adjudication; this code reports `coverage_state` and never writes Atom Markdown.

All five frozen activations are `TRUE`: no predecessor is required to compute these modules. `direct` bypasses only outer activation. Neither policy retries an explicit self `no_result`, clears scientific stops, fills missing inputs, or executes a source-prohibited internal branch. The copied native parser uses strong Kleene logic; absent evidence remains unknown, and a boundary can satisfy `done` only as a realized, accepted scientific result. Local shared construction is explicitly labeled and never impersonates an accepted predecessor Atom. Unactivated executions read no person rows and run no models; they export every requested cell with `not_run`, blank numerical fields, and the activation reason. These rows keep the frozen figure/manuscript interfaces usable without inventing a scientific boundary or silently retrying terminal evidence.

Logical inputs and field provenance

The full frozen mapping is `docs/protocol/data-contract.md`, with exact official schema locators in `docs/replication/schema-map.csv`. Those 19 local FSV documentation pages were checked again during engineering. Their labels do not establish access.

Logical input/fields	Source and candidate provenance	Required limits
cohort: pnr,start_age,stop_age,death_age; all seven incident-age columns in construction. EXCLUSION	R73-77/258-299; BEF PNR/FOED_DAG/KOEN VNDS HAEND_DATO/INDUD_KODE DOES DODDATO; clinical/LMDB history below	Pre-exclusion One-row-per-person input; D01/D02 KODEs an empirical construction from these candidates. Missing historical linkage is not a negative exclusion.
persons: pnr,pnrm,pnrf	R45-47; BEF PNR/MOR_ID/FAR_ID are candidates	Full imported person universe for parent>30 cleaning, before cohort restriction; no duplicate person IDs. Father need not be shared for sibling selection.

Logical input/fields	Source and candidate provenance	Required limits
parity1,parity_cat,bmi,ryge	R225–230/266/311/348–373; gravdb pnr/slutdato/gruppe/bmi/ryge	MFR CPR_MODER/FOEDSELSDATO/PARITET/BM age candidates only. MFR mother-child rows are not automatically pregnancy events. D06.
year_start,year_cat,fyear	R302–305/342–344/819	Source decimal calendar reconstruction and birth year; D09 horizon/calendar conflict is not repaired by extra splitting.
fland1/fland_mode,uddniv,indiv,indiv2,civst	R58165/238–255/346–380/444–448; exact source aliases	BFB deigin/pl/age is not birth country; UDDA/UDDF HFAUDD needs uddniv crosswalk; CIV CIVST/CIV_VFRA needs source conversion; IND PERIND- KIALT_13/ALDER_ULT_INK needs source pooled scope and reference year. D07. source_demographic_modes extracts R's mapping without resolving P's mode wording.
route,regimen	R384–427; derived strict cumulative dose maxima	0 never; routel oral/2 transdermal/3 other; regimen1 estrogen/2 cyclic/3 continuous. R394 puts nonmissing source forms not matching <code>trans\ oral</code> into the other-dose total, including forms outside the R94–100 map. Ties remain missing D10; no cross-person carry. Missing source definitions remain unavailable inputs.

They do not assert an empirical upstream mapping. Episode construction stops at future-purchase aggregation, covering progestin intervals, and competing matches; it preserves the uniquely determined single-episode cases. This intentionally does not select a source-silent ordering or implement a speculative repair.

Atom entry index

Every command below uses the full `run.py` syntax shown above, substituting the specified `--atom`. Every entry is exercised by `run_tests.py --output ABS_PATH`. All outputs are written into the caller's Atom output directory, never the World.

atom-000001 — Population mortality

Command: `run.py --atom atom-000001 --world WORLD --inputs MANIFEST --output OUTPUT`. Implementing functions: `Analysis.models`, `population_descriptions`, `flow_output`, `characteristics`, `blocked_displays` in `analysis.py`; `prepare_target/Engine.fit` in `models.py`; native calls in `engine.R`; shared kernels above. Inputs: cohort, intervals and relevant optional columns from the provenance table. The age55 landmark uses `tstart==55`, whereas flow, uptake, and missingness use the last `tstop` row. Follow-up sums each woman's interval lengths; it is not age at death. Cohort exclusion union is counted separately from its seven overlapping reasons.

Implemented targets: binary `m01`; duration `m11`; contraindication `m21` and preceding contraindication percentage; cessation `m41`; source-code calendar `m111/m121`, adjusted `m112/m122` without `year_cat`; surgery-adjusted `m132` (shared main crude); BMI-observed `m161/m162` with BMI30 cut; smoking-observed **summary only**; complete source descriptive roster and all Table1 cells. Published adjusted `m02/m12/m22/m42` remain D08 stops, smoking HR remains D13, S2's29 rows remain D19, PH14 panels remain D08/D18. Shared D01–D07/D09 gates remain binding. No extra trend or PH corrective model. Duration `cox.zph R480` is extracted in the engine but cannot run as an authorized empirical `m12` until D08 is resolved; no extra figure is created.

Outputs: common files below plus `flow.csv` (17 `node,n,status,reason` rows; women), `characteristics.csv` (`characteristic,category,exposure,n,percent,mean,sd,status,reason;counts/percent or parity mean/sample SD`), `products.csv` (`atc,regimen,form,percent,status,reason;29 rows`), `schoenfeld_{points,curves,panels,ticks}.csv` with the exact frozen figure schemas. Unavailable residuals/curves/ticks are empty tables with headers, and all14 panels carry reasons. Tests cover arithmetic counts/PY/rates, landmark/final timing, exclusion union, censoring, partial covariates, all cells, native survival, independent Efron inference, and figure exports.

atom-000002 — Sibling mortality

Command: `run.py --atom atom-000002 --world WORLD --inputs MANIFEST --output OUTPUT`. Implementation: `construction.clean_parents/siblings`, `models.prepare_target`, `Analysis.models/sibling_descriptions`. Inputs: original interval history and full-universe `persons`; parent cleaning must precede

eligible-sister selection. Retain all sisters in maternal families with eventual discordance, then retain their full contemporaneous exposure histories. Outputs: `t2_sibling` never/past-present summary/crude/adjusted cells; `sibling_population` `n` and `share_pct` in summaries, common coverage/adjudication. Count/PY/rate are executable independently of model attribution. Crude `m31` is stopped at D11; adjusted `m32`'s extracted formula includes `strata(family_id)` and is stopped at D08. The source has model-based covariance, not family-clustered sandwich inference. Tests include distinct fathers, missing mothers, `parent>30` in the full universe, nonmultiplying joins, eventual discordance, and a separate native stratified-Cox numerical oracle (not a repaired sibling result).

atom-000003 — Surgical mortality

Command: `run.py --atom atom-000003 --world WORLD --inputs MANIFEST --output OUTPUT.` Implementation: `models.prepare_target`, `Analysis.models/surgery_descriptions`, clinical/surgery kernels. Inputs: intervals, source surgery dates/indicators, breastgyn status, `ht_sum`. Hyst-any and BSO45–54 postoperative risk sets retain their source cancer censoring, separately; memberships can overlap. Outputs: five source blocks with 0/<5/ 5-year rows; BSO death `n` and user/nonuser death-age `q25/median/q75` in years. Crude `m71/m81` are implemented. Their adjusted targets retain D08. Older BSO and both code age-hyst targets retain the exact D12 stops; no speculative postoperative filter is run for those targets. D05 second-unilateral ambiguity stays visible. Tests retain younger outcomes when older targets stop, check event timing and BSO55 boundary, and distinguish death count from population `N`.

atom-000004 — Treatment heterogeneity

Command: `run.py --atom atom-000004 --world WORLD --inputs MANIFEST --output OUTPUT.` Implementation: `exposure_intervals/initiation`, `prepare_target`, `Analysis.models`, and `Engine.fit`. Inputs: eligible intervals, source-resolved episode DDD and doseform/regimen or the corresponding logical cumulative categories. Outputs: `t2_route`, `t2_regimen`, `t2_initiation`, each source comparator plus three exposure rows and summary/crude/adjusted components. Crude `m51/m61/m101` are implemented; adjusted attribution retains D08. Positive cumulative ties remain D10 and cannot erase separately determined initiation age; no final route is projected backwards. Initiation follows retained R618, with 52/57 boundaries; overwritten R616 is not an extra sensitivity. Tests cover multiple starts, ties, unknown dosage forms, persistent ever-use, duration epsilon/upper bounds, and independent component dispositions.

atom-000005 — Cause-specific mortality

Command: `run.py --atom atom-000005 --world WORLD --inputs MANIFEST --output OUTPUT.` Implementation: `cause_classification/cause_intervals`, `prepare_target`, `Analysis.models/cause_descriptions`, native Cox engine.

Inputs: independently resolved shorter-horizon `cause_intervals` and unique primary-cause records. Other classified deaths censor each cause-specific fit with `event0`; no risk is subtracted or competing-risk cumulative-incidence estimator substituted. Outputs: cardiovascular/cancer/other t3 rows, each 0/<5/ 5 years and three components, plus cause percentages with the common classified-death denominator. Crude m201/m211/m221 are implemented. D09/D15 remain horizon/category gates; published adjusted attribution remains D08, and adjusted other-cause remains the absolute D14 stop (m92/m222 cannot be chosen by the adapter). Tests check cause-code12 is not silently added to cancer, conflicting/missing/duplicate causes, competing death events, partial absence, and all output rows.

Outputs, inference, and portable evidence

`estimates.csv` has exactly: `target_id,exposure,component,status,reason,source_locator,n,events`. Summaries count unique women where requested; Cox's observation count means intervals and is separately recorded in `model-audit.json`. HR is dimensionless; time is years and rates are deaths per10000 person-years. Machine CSVs retain precision; `interfaces.present_estimates` formats source HR/CI to2 decimals, rates to1, person-years to0 without changing the scientific table. Never derive P from a rounded CI. Reference1 has no fabricated CI; failed coefficients stay blank. Source count<3 suppresses the group summary and its recoverable dependent rate/PY. Future disclosure control remains separate from this source rule.

`summaries.csv`: `target_id,group,statistic,value,unit,status,reason`. Every requested scalar has its own row. `coverage.csv`: `target_id,component,status,reason,source_locator,dependency`; exposure identity is retained inside model components. `adjudication.md` reports coverage, exact missing cells and interpretation limits. `source-comparison.csv` has an empty declared interface because neither synthetic results nor a governed boundary are empirical publication comparisons. Published values are never copied into outputs. `model-audit.json` records only aggregate native fit metadata, software versions, formula, warnings and aliases. No person rows/IDs or residual arrays are written into result JSON.

`result.json` separates native registration/material identity from the mutable World realization. `registration-lock.json` is a copy of the immutable source manifest; it is a content integrity anchor, not a new scientific registry. Native digest serialization is copied from `tools/prereg-lane/src/prereg_lane/native.py`. Whole Atom result bodies are not frozen material hashes. Realized body/current Molecule wording may change without invalidating faithful predecessor evidence; changed registration, membership, frozen materials, bound input contents, or corrupt aggregate/result bytes are rejected. Input identity hashes logical origin/schema/content/dictionaries, never physical runtime paths. A copied code directory can verify and read a distinct read-only World and emit results into a separate action directory; no repository import or machine source-data mount is required.

External controller trust and launch contract

`launcher.py` is a standard-library-only bootstrap to be installed by the host in its protected controller area, outside the mutable action bundle. It checks the complete action execution-directory digest **before importing `run.py` or any of its dependencies**. The digest covers every delivered file, including the registration lock, tests, README and R code, under portable relative names. Symlinks and bytecode in that directory are rejected. An estimator, source gate, verifier or lock mutation therefore rejects execution before creating outputs. `verify_world` also checks this anchor for in-process callers; a co-located verifier alone is not the security boundary.

The host independently supplies `PREREG_TRUST_FILE` (absolute registry path) and `PREREG_TRUST_SHA256` (the approved SHA256 of those registry bytes) through its protected launch settings. They must not come from the action's manifest, caller-selected environment or a hash freshly computed from unreviewed code. The registry has this operational format, separate from the scientific inputs:

```
{
  "implementation_identity": "REVIEWED_EXECUTION_DIRECTORY_DIGEST",
  "approved_implementations": ["REVIEWED_EXECUTION_DIRECTORY_DIGEST"],
  "accepted_results": {
    "HOST_ACCEPTED_RESULT_ID": {
      "path": "/HOST_SUPPLIED_ACCEPTED_RESULT_PATH/result.json",
      "sha256": "HOST_ACCEPTED_RESULT_BYTE_SHA256",
      "implementation_identity": "REVIEWED_EXECUTION_DIRECTORY_DIGEST"
    }
  }
}
```

An empty `accepted_results` object is sufficient for this all-TRUE schedule. The host can explicitly retain earlier reviewed implementation digests for faithful predecessor evidence; an unapproved digest never becomes eligible because a result is otherwise self-consistent. Host adjudication owns the accepted result's registered state and digest. Explicit accepted `no_result` still prevents retries under both policies. Unsigned/caller-forged evidence rejects the invocation; it is not converted into an operational terminal or an activation decision.

The directory digest is SHA256(canonical JSON of {relative filename: file SHA256}), using sorted keys, UTF-8, `ensure_ascii=False` and separators `,/:`, as implemented in `launcher.bundle_identity`. The reviewed snapshot's digest is recorded in the engineering report for host review; that report does not authenticate itself. Once the protected bootstrap and registry are provisioned:

```
/home/shaoerzhuo/anaconda3/bin/python -I -B "$CONTROLLER_LAUNCHER" \
--bundle "$ACTION_EXECUTION" --atom atom-000001 --world "$WORLD" \
--inputs "$INPUT_MANIFEST" --output "$ACTION/results/atom-000001" \
--dispatch-policy activation_respecting
```

All variables are explicit host paths. `-I` excludes caller Python startup paths; the host must prevent concurrent mutation of the verified action snapshot and protect its registry, pin and bootstrap. Host deployment was not performed by this author, who cannot write controller state. Missing trust configuration fails closed. The synthetic

suite executes a separate bootstrap with an external test registry, tests copied-code mutations including `run.py` and the verifier, and checks registry substitution under an unchanged host pin. Those credentials are labeled test-only and are never real predecessor acceptance.

R inference calls `coxph(Surv(tstart,tstop,event) ~ source_terms)` with Efron ties and model-based variance, and uses native `summary(..., conf.int=.95)`. There are no weights, matching, clustered sandwich, frailties, extra interactions, modern DiD/RD approximations or alternative survival models in this source. `na.fail` prevents an unregistered complete-case fallback: synthetic tests supply resolved covariates, while an empirically consequential missing-value default remains a source prerequisite. Factor labels/reference order are explicit source labels. Empty/no-event/no-contrast groups, aliased designs and inference warnings retain non-estimation dispositions. Unexpected engine failures produce `no_result`, even when successful sibling outputs remain in the artifact.

Absent factor levels and constant adjustment covariates retain aliased nuisance coefficients while independently identifiable contrasts survive. Native Cox handles both all-zero absent dummies and an all-one constant dummy without a refit; the same holds for constant numeric source indicators. The audit records numeric constants in `constant_numeric_terms` and in the compatibility field `constant_factor_terms`. Exposure constants still have no contrast; `family_id` is a stratum and is never included in the allowed alias set. Other rank deficiency involving observed columns stops the consuming fit without refitting a smaller model. If local native interval construction fails, its consumers carry the exact boundary or operational disposition. Independently supplied cause histories do not require that construction. An episode beginning inside an unsplit interval is rejected as a missing `htstart` change, rather than silently disappearing in a left join.

Figure and manuscript interface checks

```
/home/shaoerzhuo/anaconda3/bin/python -B docs/execution/render_figures.py \  
--world "$WORLD" --results "$ACTION/results/atom-000001" \  
--output "$ACTION/figure-checks" --consumption preview
```

This explicitly requests a noncanonical preview from the unchanged `figure1.py/figure2.py` renderers at 130.7mm. Supply `--style-dir` only when the actual host fonts exist. `interfaces.select_slots` selects every 292 fixed manuscript binding into existing output tables/prose; it is a read-only interface checker, not a manuscript updater. Tests render generated outputs and complete/partial/all-missing/unactivated interfaces, check PDF width, and reject numerical values attached to unavailable inputs. Synthetic panel arrays used to test figure2 are arbitrary display fixtures, not fitted PH results. No synthetic preview fills a canonical Atom or manuscript slot. Both this selector and the figure adapter use `provenance.read_result` to check the result digest, frozen material, Atom identity, approved implementation, required file membership and all aggregate hashes before consuming a file. Their default `consumption="canonical"` additionally requires the exact result byte hash and implementation to be bound in the externally authenticated `accepted_results` registry, plus `synthetic=false` and `accepted=true`.

Rehashing edited CSVs and resealing a result envelope cannot create this acceptance. Result copies with identical bytes remain valid after relocation. The host retains responsibility for scientific adjudication and disclosure.

For synthetic or unaccepted interface checks use `select_slots(world, results_root, consumption="preview")` and `render(world, atom1_results, output, consumption="preview")`, or the CLI flag above. Preview mode never licenses insertion into a manuscript or canonical Atom. The figure adapter writes `PREVIEW-ONLY.txt`, retains `SYNTHETIC-ONLY.txt` when appropriate, and writes `figure-provenance.json` with the consumed result byte hash, material/input/code identities, flags and consumption mode.

In both modes the slot selector requires all contributing Atoms to share one `input_identity` and one `synthetic` flag. Every returned row includes those values, `accepted`, `consumption` and its `result_sha256`, so CSV export retains provenance; shared identity/flag/mode are also DataFrame attributes. All 292 bindings remain unchanged. Mixed partial/not_run dispositions from the same inputs are valid; combining artifacts from different input manifests is rejected.

Remaining limits and the exact author test record are in `../engineering/implementation-report.md`, including responses to every finding from both prior rounds. Independent round-three reviewer execution remains the controller's subsequent stage; no agents were spawned by this author.

Engineering supplement

Author implementation report — round 3

Date: 2026-09-06. Scope: the five existing canonical Atoms, atom-000001 through atom-000005. No membership, registration, source audit, manuscript, or figure template was changed. This report describes executable pre-data engineering, not an empirical replication or an accepted Atom result.

Round-three changes and responses to every round-two finding

The supplied round-two code was retained and revised only in execution/ and this report. Both prior Opus and Codex reviews were read. Round two's Opus verdict was accept with F5/F6 fixes; Codex's was revise with C2-01–C2-03. The earlier round-one responses below remain part of this report. No source rule, native activation, Atom membership, manuscript binding, or fixed figure template changed in this round.

Finding	Implemented response and why it preserves meaning	Regression evidence
Codex C2-01: canonical consumers accept caller-resealed results	<code>provenance.read_result</code> defaults to canonical consumption and requires the exact result byte hash plus implementation ID in the protected external accepted-results registry, in addition to all existing content checks. Synthetic/unaccepted artifacts are never canonical. Explicit <code>consumption="preview"</code> preserves noncanonical testing with visible flags/markers. Paths are not acceptance identities: copied identical bytes remain valid. No scientific estimate or decision rule changes.	<code>test_canonical_consumers_reject_resealed_re</code> starts with genuinely generated governed boundary artifacts, exercises simulated external acceptance, then edits an HR and flow count to 999, updates both CSV hashes and reseals the result. Public integrity succeeds in preview mode; both canonical consumers reject before rendering. Positive canonical checks use only nonnumerical governed boundaries, never promoted synthetic estimates.
Codex C2-02: missing declared dataset hash	<code>load_manifest</code> requires a 64-hex SHA256 for each declared dataset before dispatch or any person-row access; valid hex is normalized to lowercase. This checks a declaration, not runtime access or content. Omitted optional datasets remain absent; the empty governed manifest still yields its honest frozen boundary.	<code>test_declared_datasets_require_hashes_befor</code> rejects absent, empty, short, long, nonhex and nonstring hashes for both purposes before output. It accepts a well-formed declaration without opening its nonexistent path and verifies the empty governed case. Actual hash comparison in the existing frame-loader test remains.

Finding	Implemented response and why it preserves meaning	Regression evidence
Codex C2-03: source indicator domains and calendar continuity	<code>require_binary</code> checks consumed R222/R309–330/427 indicators and R228 smoking values. Fits validate only actual source covariates; filters/descriptions validate only fields they use. R342–344 calendar construction and R633/R647 preparation reject unexplained internal gaps before their arithmetic/selection. No gaps are filled, no source clock is replaced, and unrelated risk-time outputs remain independently computable.	<code>test_source_binary_domains_are_checked_only_by_</code> exercises invalid disease/surgery/cessation/smoking values, missing and nonfinite values, every relevant filter, and Table1; invalid dm stops adjustment and its descriptor while crude/summary siblings survive. <code>test_calendar_rejects_gaps_without_erasing_nonc</code> rejects (45,46],(50,51] for calendar processing and preserves its two supplied risk-years elsewhere; contiguous shuffled splits give2000/2005.
Opus F5: constant numeric nuisance covariates	Native Cox retains the full source formula. Constant numeric nuisance columns join the allowed all-zero/all-one alias set; <code>family_id</code> and exposure are excluded from this allowance. Audit adds <code>constant_numeric_terms</code> and preserves the existing <code>constant_factor_terms</code> field for compatibility. No reduced model is exported and no covariance changes.	<code>test_constant_numeric_covariates_native_oracle_</code> checks dm=0 and statin=1 against independent direct R fits omitting only the constant term as numerical oracles. Exposure beta/model-based SE agree to seven decimal places. Varying numeric and factor collinearity still stop, and <code>family_id</code> never appears as an allowed alias.

Finding	Implemented response and why it preserves meaning	Regression evidence
Opus F6: mixed input identities in slots	<code>select_slots</code> requires one shared input identity and one synthetic flag in both modes. Every returned row carries identity, flags, consumption mode and result-byte hash; shared metadata also appears in DataFrame attributes. Fixed slot identities, fields and numerical values are unchanged.	<code>test_slot_consumption_keeps_one_input_identity</code> rejects a mixed synthetic/governed tree and mismatched flags even after public resealing. Existing partial, governed-boundary and <code>not_run</code> checks still select 292 rows and verify their shared provenance.

Opus’s request to preserve previews and Codex’s request for authenticated canonical consumption are addressed together: preview checks remain available by explicit choice; they carry no manuscript-use permission. Figure previews write `PREVIEW-ONLY.txt` and `figure-provenance.json`, retaining the synthetic marker as applicable. The external controller contract already used for predecessors supplies canonical acceptance; no new scientific registry or host-approval claim is introduced.

The round-three author did not spawn agents or initiate new independent reviews. Adversarial Opus and Codex Sol xhigh checks remain the controller’s subsequent stage, as required by this role’s write and delegation restrictions.

Delivered coverage and its principal limit

The implementation contains actual construction functions, native R counting process and Cox calls, component exporters, portable evidence verification, and synthetic numerical/interface tests. The entry index and complete input/output contract are in [execution/README.md](#). The CLI is [run.py](#); the reproducible test entry is [run_tests.py](#).

No empirical cohort is authorized by the frozen source. D01 and D02 stop every cohort-derived target at unresolved entry/age and migration/censoring rules. Governed execution therefore emits the precise registered boundaries before opening any person-level CSV. A supplied derived table, a synthetic flag, or resolution metadata cannot amend those rules. This limitation follows [replication/deviations.md](#), D01/D02, every Atom’s COMMON POPULATION / TIME paragraph, and [governance.md](#) paragraph 2, which requires a documented pre-execution amendment for scientific source clarification.

The independently supported lower stages are implemented and executed on explicitly synthetic inputs. Hard target conflicts remain closed even there. There is

no placeholder estimate, invented standard error, alternative Cox model, or automatic missing-to-one repair. This is not an end-to-end empirical replication engine that can presently be pointed at Danish data and produce the paper. The distinction is material, not merely an access disclaimer.

The frozen catalog contains 21 estimate blocks, 32 characteristic rows, 29 product rows, 12 summary targets, one flow, one diagnostic and six context items. All requested non-context components have exact output identities. Context items do not become extra statistical analyses.

Atom	Implemented supported execution	Source-stopped siblings
atom-000001	Flow/exclusion union; last-interval and age55 descriptors; Table1; binary m01, duration m11, contraindication m21, cessation m41; code calendar m111/m112 and m121/m122; surgery-adjusted m132 with shared crude; BMI m161/m162; smoking-complete summaries	Published adjusted m02/m12/m22/m42 D08; smoking fits/table D13; 14 PH panels D08/D18; 29 product proportions D19; all empirical consumers retain shared D01–D07 and applicable D09
atom-000002	Full-universe parent cleaning; maternal eventual-discordance selection; contemporaneous histories; population count/share and exposure-specific counts/PY/rates	Printed crude sibling fit D11, adjusted attribution D08. No extra strata were added to m31 and no full-sibling restriction was substituted
atom-000003	Hyst-any and postoperative BSO45–54 summaries and crude m71/m81; breast/gynaecological cancer censoring; BSO death count and conditional death-age quartiles	Adjusted D08; all older-BSO and both code age-hyst outputs D12; historical/laterality prerequisites D05

Atom	Implemented supported execution	Source-stopped siblings
atom-000004	Cumulative dose/strict dominant-type classification, including R394 other forms; route m51, regimen m61, retained initiation-age m101; per-group summaries; independent initiation branch when route fails	Published adjusted D08; ambiguous episode construction D04; affected tied type inputs D10
atom-000005	Separate cause-horizon logical histories; known primary cause classification; other deaths as censoring; crude m201/m211/m221, summaries and cause percentages	Empirical horizon/version/missing-cause D09/D15; published adjusted D08; adjusted other-cause extraction D14

In the expanded synthetic suite, all 20 permitted fit invocations (including the shared crude export) produced their expected 35 nonreference HR cells. The five Atoms exported respectively 225, 8, 52, 36 and 30 component coverage rows. There were 214 available rows, 20 structural references and 117 source-stopped rows; all five synthetic coverage assessments were partial. These are engineering coverage counts, not study findings. No synthetic value was copied to a canonical Atom, manuscript slot, or published-reference comparison.

Authority and source checks

Read the lane AGENTS.md; the full predata-engineering, preregistration-package, paper-to-replication and applicable canonical skill instructions; required package references; all five complete Atom protocol/result regions; the complete Molecule registration; shared data-contract, identification and governance; all source intake/audit/deviation and mapping documents; the full 969-line R source; existing paper result/figure interfaces; and both prior engineering reviews. There is no **prereg/code/**. The supplied round-two execution code was read in full and revised in place, preserving its supported models and kernels. Source-native canonical parsing/digests were adapted from the repository’s **prereg_lane/native.py**; no adaptive schedule was introduced.

P, S and R retain exactly the locators defined in **replication/source-map.md**: P is the ten-page BMJ article, S the four-page supplement, and R the supplied 969-line CRLF code. Original bytes were not rewritten or redistributed. All seven local supplied

source files matched `source-provenance.yaml` during this author check. This verifies the supplied material identity, not publisher authentication of the user-supplied R file.

The 19 local FSV schema pages referenced by `schema-map.csv` were inspected for the relevant field meanings. They support candidate mappings, not runtime availability. The existing 531-table boundary remains unchanged. In particular, LMDB VOLUME is not established as DDD; LPR_DIAG links by RECNUM rather than an invented person field; procedure timestamps are distinct from admission dates; BEF place/origin is not proven birth-country equivalence; MFR fields do not establish the gravidb pregnancy construction. The README links each logical field family to its source code and candidate official fields.

No confidential Danish/APTO tables, old empirical sessions, or private measured outcomes were read. Public published values were read only as source authority; none was used to generate a current result. No live search was used in this round, so there are no new queries or URLs to report. Historical publisher 403 and unauthenticated-R limitations remain as recorded in the frozen source map; this author did not claim to resolve or re-test them.

Construction and meaning-preserving adaptations

`construction.py` implements source-age arithmetic as an extracted kernel, many-to-one/complete-link checks, source exclusions, full-universe parent rules, prescription classification/supply, invariant episode cases, 90-day cessation, ordered medication classes, clinical event dates, heart-failure timing, diagnosis-row poor health, pregnancy processing, demographic mappings and rolling joins, pooled income quantiles, calendar reconstruction, parity, duration/dominant type, initiation, sibling selection and known causes.

`engine.R` uses actual `survival::tmerge`, `survSplit` and `coxph`, including the source `tdc/cumtdc` distinctions, `age50/55/60/65/70/75` splits, Efron ties, model-based covariance and native 95% intervals/P values. It has no weights, matching, sandwich clustering, frailties or alternative survival estimator: none is specified in this source. The separate stratified numerical oracle checks mother-stratum mathematics without calling it an authorized repaired sibling analysis. Published adjusted calls are extracted but remain gated.

The mechanical adaptations preserve the supported estimand as follows:

- String person keys and explicit CSV schemas preserve identifiers. Duplicate right-side links, missing required keys, overlapping intervals, and multiple or nonterminal deaths fail visibly instead of multiplying records or treating missing history as zero. Lower-stage input names are not asserted FSV tables.
- Native R handles event-at-boundary and time-dependent changes. An episode inside an unsplit interval is rejected. No new split date or retrospective exposure assignment is invented to repair an incomplete input.
- Named source coefficients replace fragile positional recycling. Only the specified reference is one; it has no fabricated CI. Absent factor levels retain missing coefficients while identifiable contrasts survive. Rank deficiency among observed columns and inference warnings stop the consuming fit without dropping covariates or choosing a smaller model.

- `na.fail` makes incomplete scientific inputs visible rather than silently adopting complete cases. Source-explicit BMI/smoking observed-interval restrictions are implemented separately. Missing optional columns do not erase unrelated crude fits or summaries.
- R type7 quantiles and sample SD are used for source descriptors. Pooled income ranks require the supplied source record universe; repeated quantile breaks or ambiguous person-age ties stop instead of selecting a new ranking rule.
- Literal supported demographic and calendar arithmetic is callable for tests; it does not settle D07/D09. The D01 birthday and numeric-date error, D04 future dose booking, D12 retrospective surgery, D13 wrong smoking object, and D14 wrong extraction object are not repaired.
- Raw Cox precision is retained in machine outputs. Presentation formatting follows source precision without reconstructing P values from rounded CIs. Source event counts below3 and recoverable dependent group summaries are suppressed together. This is not a claim of future disclosure clearance.
- Operational failures retain `no_result` and no local state, including failure during shared construction. Known source/input absence remains a named boundary. Successful independent outputs can stay in a `no_result` artifact, but cannot turn an unfinished module into partial success.

Exact unresolved-source inventory

The following is an execution crosswalk, not a replacement or amendment of the frozen ledger. `unresolved_source/unavailable_input` statuses correspond to its `STOP_UNRESOLVED_SOURCE` / `STOP_UNAVAILABLE_INPUT` dispositions.

Frozen ID and exact locator	Consuming target and retained stop
D01: P2–3; R39–40,73–77,281–297	All empirical cohort-derived cells: birthday versus365.2, boundary inclusivity and numeric <code>as.IDate(1995-01-01)</code>
D02: P3; R51–53,72,76	All empirical follow-up/risk sets: migration order/status, six months versus180 days, open departure
D03: P2; R80–141	MHT input equivalence, pname/NAME, APK×VOLUME/DDD, pre1995 coverage; no clinical-dose dictionary invented
D04: R112–132,384–389	Future episode dose, covering/competing progestin intervals, duration/type/cessation consumers. Independently determined binary use is not rejected merely because a type is missing

Frozen ID and exact locator	Consuming target and retained stop
D05: P3/S1; R174–222	Exclusions, surgical history, clinical covariates/Table1: original cancer/unified surgery, laterality/order, admission/discharge, diagnosis/contacts
D06: R225–230,266,311,348–373	Pregnancy/parity/BMI/smoking consumers: gravdb event equivalence is unavailable; no MFR substitution
D07: R238–255,346–380,444–448; P4	Demographic descriptors/adjusted consumers: record universe, crosswalks, rolling ties, mode versus hardcoded fill
D08: P1/P3/T2/T3; S4; R461,518,756 and adjusted calls	Every published adjusted T2/T3 output and m02-based PH attribution; no alternative adjustment set estimated
D09: R302–325,342–344,633/647,740–777	Empirical calendar/cause horizons: interval-start rules and exact calendar boundaries; no extra2003 split or chosen truncation
D10: R397–405,418–426	Affected route/regimen ties; no global or repaired within-person carry. Initiation remains independent
D11: P3; R45–47,505,515/518	Crude sibling HR lacks resolved within-family attribution. Summary siblings remain executable
D12: P3; R583,670	All t2_bso_ge55 and both code_hyst age blocks; no guessed postoperative or prior-surgery restriction
D13: R714–725	Smoking crude/adjusted/composite target; retain only source smoking-complete summaries
D14: R784–785; P T2/T3	t3_other adjusted: m92 extraction is not relabeled other-cause and not replaced by m222
D15: R745–746,777; S1	Empirical cause version, missing and duplicate records. A missing cause is not automatically other mortality
D16: R463 and analogous assembly; R729/789	Named authorized coefficients and explicit references only; no failed estimate/text becomes1
D17: R946–947	Do not claim whole original script execution. Unchanged Python flow renderer consumes validated exported counts

Frozen ID and exact locator	Consuming target and retained stop
D18: S4; R465–466,480	Exact ggcoxzph transform/smoothers/version and model attribution unresolved; blank 14-panel PH interface, no substitute diagnostic or acceptance threshold
D19: S3; R430	Each of 29 S2 proportions and most-common-product prose: unresolved denominator; no DDD/person weighting invented
D20: R632–725,832–840	Retain only registered code-only sensitivities. Density and commented explorations remain context, not new experiments

Dispatch, material identity and output interfaces

All five native activations are TRUE and need no accepted predecessor. **activation_respecting** is the CLI default; **direct** bypasses only outer activation. Strong Kleene unknown/false/true and explicit no_result terminals are tested. Neither policy retries an Atom explicitly marked no_result. Unactivated artifacts retain every frozen cell as blank **not_run**, without reading person inputs or calling the engine. Local shared construction is recorded only when attempted and is never an accepted predecessor.

registration-lock.json copies the original manifest as the material anchor; its own bytes and all execution files are now externally anchored as below. Registration/material identity uses native canonical digests and frozen material hashes; mutable Atom result bodies and Molecule interpretation have a separate realization digest. Tests mutate only scratch copies of those bodies and verify that faithful earlier evidence remains valid. Protocol changes, wrong input or material binding, corrupt result/aggregate content, and invalid scientific states are rejected. Host acceptance now requires a controller registry entry authenticated by a digest supplied outside the action manifest. A result’s own accepted Boolean is insufficient. This executor never self-accepts a result or writes an Atom body.

Input identity includes logical schema/origin/content and dictionary/source metadata, excluding physical runtime paths. Absolute paths hidden in scientific metadata are rejected. Tests use copied code in a writable action directory with a separate read-only World and outputs outside it. No import from the repository source or a presumed data mount is required. Person-level R exchange files are temporary and removed; result JSON contains aggregate provenance, not raw rows.

The fixed manuscript has 292 slots. They are selected read-only against generated tables, including unactivated, partial and governed boundary dispositions. The selector and figure loader now share result-digest, material, Atom, approved-code, required-file-membership and aggregate-hash verification. No manuscript write occurs. Both unchanged figure renderers consume generated interfaces at 130.7mm. Complete,

partial, all-missing and unactivated interfaces are rendered; incompatible missing values are rejected. PH numerical panel fixtures are arbitrary renderer inputs, not fitted Schoenfeld results. Partial plots were visually inspected at manuscript width: numbered explanations remain readable, without editing templates. Both unactivated plots were visually inspected and remain readable. Helvetica Neue was unavailable; the template’s DejaVu Sans fallback is used only for local checks, with a documented host `style_dir` font binding for publication.

Response to every round-one reviewer finding

The source-fidelity findings below are Opus F1–F4 from `reviews/r01/opus.yaml`. Codex supplied titles without IDs; C1–C3 below identify those three titles in their original order. Both independent round-one verdicts were **revise**. These responses implement engineering corrections, not scientific amendments.

Finding	Response and source-preserving reason	Executed regression
Opus F1, omitted R319 cancer tmerge	Added cancer to the native tdc whitelist and Python input validation. <code>clinical_dates</code> retains the literal R198 hospital and R218–219 cancer-file predicates for the incident split date; missing historical cancer stays D05. It is only a source interval split, never a covariate or new outcome. Unknown change variables now yield <code>unavailable_input</code> , not an R crash. No 2003 split was added.	<code>test_r319_cancer_split_calendar_filter_and_negate</code> retains the 48.6 split and 15 total years; R633/R647 selections are 3.6/11.4 versus 5/10 without that split. Aligned year-category totals remain invariant. Clinical-date tests cover C44/191 exclusions, numeric cancer codes and missing cancer history.
Opus F2, terminal episode selected by last start	<code>cessation</code> now marks the largest episode end , then takes the earliest retained gap/terminal stop, exactly R136. Tied maximum ends give the same returned age; grace equality is not a strict gap. No new cessation rule or reentry.	<code>test_r136_nested_episodes_and_pipeline_oracle</code> compares against literal R136 in <code>data.table</code> . The nested toy yields 55.246; the <code>prescription→episode→source-age</code> pipeline yields 56.182. The prior ordinary-gap test remains.

Finding	Response and source-preserving reason	Executed regression
Opus F3, unmapped forms incorrectly stopped	Removed the added route nullification. R394 contributes every nonmissing form not matching trans/oral to other dose. Raw strings remain unchanged at R94–100; this is not a clinical crosswalk. Genuine cumulative ties still stop D10; missing dose-form definitions remain unavailable.	<code>test_r394_other_form_preserves_separate_mis</code> gives route 3 for SUSP, retains unexposed 0, and rejects missing form. The existing oral/transdermal and added other/transdermal ties remain missing.
Opus F4, constant adjustment factor stops fit	Removed the pre-fit singleton-factor guard. Native Cox retains the full formula and handles aliased all-zero/all-one nuisance columns; the audit names absent and constant terms. Other collinearity and no exposure contrast still stop. No smaller model is exported.	<code>test_constant_adjustment_factor_native_orac</code> tests both observed levels of constant fland_mode with varying other covariates. Exposure beta and model-based SE match independent direct R oracle calls omitting only the constant term. Existing collinearity/crude-retention tests pass.
Codex C1, caller JSON impersonates acceptance	Manifest predecessors now carry only controller result IDs. <code>verify_evidence</code> resolves an externally authenticated accepted-results registry, checks its result byte hash and approved implementation, then checks material/input/state/output bindings. Caller paths, invented IDs and a self-asserted accepted Boolean are insufficient.	<code>test_unsigned_forged_acceptance_cannot_cont</code> rejects forged boundary/no_result records and unapproved implementation IDs before outputs. Direct dispatch without such forged evidence remains TRUE. <code>test_cli_explicit_no_result_never_retries</code> authenticates a test terminal externally and confirms both policies refrain from retrying.

Finding	Response and source-preserving reason	Executed regression
Codex C2, executable identity recorded but unanchored	Added standalone <code>launcher.py</code> : a protected external copy validates the whole execution-directory digest against host-pinned registry bytes before importing <code>run.py</code> . The lock, estimators, verifiers, tests and README all contribute by relative name. Internal verification also fails closed without the external anchor. See deployment limit below.	<code>test_external_launcher_rejects_code_lock_and_tr</code> changes <code>models.py</code> , <code>engine.R</code> , <code>registration-lock.json</code> , <code>native.py</code> , <code>launcher.py</code> and <code>run.py</code> in action copies. Every launch rejects before imports/outputs. Forged registry bytes and absent host pin reject too. The separate read-only relocated World test succeeds through the external launcher.
Codex C3, slots ignore sealed aggregates	Factored <code>read_result</code> into <code>provenance.py</code> . Slots and figures both check result digest, material and Atom identity, approved implementation, required file membership and all aggregate hashes, including Markdown. This is integrity checking, separate from host acceptance for empirical use.	<code>test_slots_enforce_all_aggregate_envelopes_and_l</code> rejects a primary crude HR changed to 999, altered adjudication prose and a resealed envelope omitting the selected CSV. All 292 original slots select with valid partial, not <code>_run</code> and all-five governed boundary artifacts.

Opus explicitly withdrew the broader `year_cat` claim in F1; no such repair was made. Its adjacent comments about the unbound `n` field and dependent suppression were not findings and remain as registered. Neither reviewer requested changing D01/D02 or any other frozen source stop; all remain binding.

Controller deployment boundary

The author is authorized to deliver code and documentation, not to install controller settings, approve a bundle, or create real accepted evidence. The protected launch adapter and its external-registry format are complete and synthetically exercised. **Actual controller deployment/approval is not claimed.** The host must install the reviewed bootstrap outside the mutable action copy, independently pin the

reviewed execution digest and trust-registry SHA256, and prevent concurrent modification of the verified snapshot. A caller-provided environment or a hash computed from an unreviewed directory is not an approval.

The contract is documented in `execution/README.md`. Without `PREREG_TRUST_FILE/PREREG_TRUST_SHA256`, execution rejects before outputs; `-help` still works. The tests create a separate `SYNTHETIC-CONTROLLER` area with labeled test credentials, not production acceptance. Earlier implementation digests can remain explicitly approved by the host for faithful predecessor evidence. Mutable Atom result bodies and Molecule interpretation never enter material identity, while protocol and membership changes still reject.

Historical round-two commands and executed verification

This section preserves the supplied round-two test record and digest; these are not identifiers for the revised round-three code. Current verification is recorded separately below.

No software was installed or shared environment changed. Rechecked versions: Python 3.12.4; NumPy 1.26.4; pandas 2.2.3; SciPy 1.15.2; PyYAML 6.0.2; Matplotlib 3.10.0; statsmodels 0.14.4; pypdf 6.11.0. Actual synthetic native execution used R 4.1.2, survival 3.8.3 and jsonlite 2.0.0. `data.table` 1.18.2.1 executes the independent literal-R136 oracle, not the adapter.

The frozen R 4.4.1/survival 3.7-0/data.table 1.16.2 environment remains unavailable. Non-synthetic engine use enforces the R/survival pin. `survminer` is unavailable and its exact source version remains unresolved; no substitute PH inference was performed. Publication Helvetica Neue resources remain absent; unchanged figure templates use their documented DejaVu Sans test fallback.

Reproduction, with an explicitly supplied fresh absolute scratch directory:

```
/home/shaoerzhuo/anaconda3/bin/python -B \  
  /home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-code-2026090 \  
  ↪ 6/hormone-therapy/v002/world/prereg/execution/run_tests.py  
  ↪ \  
--output /home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod \  
  ↪ e-20260906/hormone-therapy/scratch/r02/author/NEW_TEST_DIRECTORY
```

Round-two runs are only under `scratch/r02/author`:

- **test-01:** 32 tests, zero failures/errors, 82.598 seconds, World unchanged. This includes all seven reviewer regressions, the prior 25 checks, real native fits for every supported Atom path, all 292 slot bindings, and unchanged figure renderers in complete/partial/missing/not_run conditions.
- After test-01, added literal source all-cancer incident-date construction and its cases to the existing clinical-date test, and completed execution README trust documentation. Final-snapshot verification is recorded below.
- **final-validation: 32 tests passed in 84.033 seconds, zero failures/errors, World unchanged.** All 20 supported model invocations produced the expected 35 nonreference HR cells. All five synthetic modules remain partial with 214 available rows, 20 structural references and 117 explicit source stops.

All five governed modules emitted boundary dispositions with absent numerics. Test log, report, slot checks, synthetic inputs and outputs are under `scratch/r02/author/final-validation/`; the aggregate engineering check is `scratch/r02/author/final-validation-summary.json`.

The historical round-two final execution-directory digest is 58318223c36ebf3a0c91ef9c018f5606ec63044b6f29fcb799abfd00396f8844. The separately deployable launcher file SHA256 is 944caf0cd41cbdf71a3d596a6f72f40436ad24e3b089425b6a35dc959. These identify the tested source snapshot for controller review, not approval. The engineering report is outside the execution-directory digest, avoiding a self-referential hash. `final-world-integrity.json` in author scratch records zero altered original files and the exact authorized added-file inventory.

Both partial and not_run figures were visually inspected via 110 dpi rasters at the registered 130.7 mm width. Labels, panel order, nearby endpoints and complete missing-reason notes remain readable without clipping or overlap. All ten final PDF exports are byte-identical to the corresponding test-01 exports; fixed templates and their interfaces were not edited. Native parser/evaluator/digest functions were also compared by AST to the repository originals: all match.

The suite includes independent Efron score and PHReg coefficient/SE/95%CI/P-value oracles, an independent stratified-Cox oracle, literal R136 and direct native constant-factor oracles, hand counts/person-time/rates/type7 quartiles, and native tmerge boundary expectations. Empty/singleton/no-event/rank-deficient groups, partial optional inputs, multiple treatments, cumulative ties, competing deaths, source gates, tri-valued dispatch, both policies, absent and authenticated terminal evidence, corruption, protocol tampering and relocation are exercised. Seed 20260906 is only for synthetic fixtures.

Integrity: all 79 original World files and all seven supplied source hashes match their immutable originals. Only 15 execution files and this engineering report are added relative to the immutable original. Source checks and all generated data/results/previews remain in author scratch. No Atom, manuscript slot, source audit, template, controller state or remote QGPU was written.

Current round-three commands and executed verification

All round-three generated inputs, results, previews, logs and test credentials are under `scratch/r03/author/` in the task workspace. Reproduction uses a fresh absolute output directory:

```
/home/shaoerzhuo/anaconda3/bin/python -B \
/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-code-2026090
↪ 6/hormone-therapy/v003/world/prereg/execution/run_tests.py
↪ \
--output /home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
↪ e-20260906/hormone-therapy/scratch/r03/author/NEW_TEST_DIRECTORY
```

- **baseline-tests:** supplied round-two code, 32 tests passed in 82.779 seconds, zero failures/errors and World unchanged.

- **revision-tests:** revised code, 38 tests passed in 101.126 seconds, zero failures/errors and World unchanged.
- **final-tests:** after making invalid-value fixtures explicitly floating point to avoid pandas assignment warnings, **38 tests passed in 100.659 seconds**, zero failures/errors and World unchanged. This is the final execution snapshot. **final-tests/test-report.json** and **final-tests/test-log.txt** contain the actual machine verdict and individual tests. **pinned_engine_validated** remains false.

The versions listed in the historical section were rechecked and are unchanged. No engine, package or shared environment was installed. Native Cox calculations were executed numerically, not merely parsed. The expanded suite preserves the independent Efron/PHReg, stratified, counting-process, literal source-R and hand oracles, and adds independent native constant-numeric comparisons. Both dispatch policies, missing/terminal predecessors, component-level input failures, mutable realized World bodies, protocol/code/aggregate tampering, and relocated read-only Worlds with writable action copies are exercised.

audit_delivery.py in author scratch was executed with the same Python. Its **delivery-audit.json** records all 79 immutable original files unchanged, all seven supplied source hashes matching, 92 exact schema-field records on 19 local pages, 102 frozen targets, and all 292 unchanged manuscript bindings. The native parser/evaluator/digest functions match the repository authority by AST. Only the 15 execution files and this report are added relative to the original. Every Atom's scientific CSVs equal the round-three baseline, including numerical values and component dispositions; audit/provenance metadata is intentionally expanded. All five governed outputs contain no numerical estimate or summary.

The tested round-three execution-directory digest is 92459bc6ebbe979f09f0775db64aa5841a1f8f926f8e. The separately deployable launcher SHA256 remains 944caf0cd41cbdf71a3d596a6f72f40436ad24e3b08942. These are review identities, not controller acceptance. Only this report was updated after final verification; it is outside the execution digest.

Figure interface investigation and remaining presentation defect

All six complete/partial/missing interface-fixture PDFs and four generated or unactivated PDFs are byte-identical to baseline. Two additional governed-boundary PDFs exercise canonical consumption using explicitly simulated host acceptance. All twelve final PDFs equal the revision-test exports. Partial Figure1/Figure2, unactivated Figure1, and the actual generated source-stopped Figure2 were visually inspected at 110 dpi. The diagnostic arrays in complete/partial display fixtures remain arbitrary transport tests, not valid fitted PH diagnostics.

The visual check found a **pre-existing Figure2 note-clipping defect**, despite successful field/disposition/width smoke tests. Author-scratch **audit_figure_bounds.py** independently invokes **pdftotext -bbox** on all twelve final PDFs. **figure-bounds-audit.json** records a page width of 370.488189 points (130.7 mm) and these exceptions:

- `generated-figures/figure2.pdf`: the word `exact` in the `year_cat` missing-reason note extends to `x=371.450969`, `0.962780` points beyond the page.
- `canonical-boundary-figures/figure2.pdf`: two boundary-note words extend to `x=373.957094` and `x=370.555094`, respectively `3.468905` and `0.066905` points beyond the page. No out-of-page words were found in the other ten PDFs.

The cause is the fixed character-count estimate in `prereg/paper/figures/figure_common.py:67-74`, combined with the `Figure2` note placement at `figure2.py:104-105`; actual DejaVu Sans glyph widths can exceed that estimate. This finding does not establish behavior with the unavailable publication fonts. The frozen templates are outside author write scope, so no template, note wording, data value or plot width was changed to conceal it. Full reasons remain in the CSVs. These two PDFs are not certified as unclipped publication exports. Template repair remains an explicit presentation limit for the controller; it does not alter any scientific target or stop disposition.

Review status and limitations

This round-three author response addresses both rounds of prior reviews; it does not impersonate a new Opus or Codex Sol xhigh review or claim their acceptance. The invocation forbids spawning agents and restricts the role to author outputs. Independent round-three checks remain a subsequent controller stage.

All empirical numerical targets remain unavailable under the frozen D01/D02 cohort rules, with additional component limits tabulated above. Source comparison stays empty. There was no confidential input access, measured study outcome, empirical execution, pinned-engine validation, real controller acceptance, extraction-equivalence verification or disclosure clearance. Passing synthetic tests is neither empirical replication, causal validation nor proof of data availability.

Reference implementation review

Engineering status: REVISE — retained at handoff. Review rounds: 3. Revision slots used: 2 (initial draft excluded).

User-directed bounded handoff: no more than two author revisions. Proceeding to formatting is not reviewer acceptance. Failed/missing reviewer calls have no scientific verdict; original records remain available below.

This augmentation used schemas, documentation and synthetic fixtures only. Code review and package compilation do not establish scientific validity, empirical replication, data availability or independent acceptance of the original science. Original registrations, manuscript sources and scientific review remain unchanged. Unresolved source definitions remain unresolved; runnable components and blocked targets must be distinguished in `prereg/execution/README.md`.

The host checks prose Atom coverage, Python syntax and synthetic commands. These are engineering checks, not a scientific verdict.

Controller checks and review dispositions

```
{
  "round": 3,
  "accepted": false,
  "host": {
    "status": "pass",
    "findings": [],
    "processes": [
      {
        "label": "r03-host-run",
        "command": [
          "/home/shaoerzhuo/anaconda3/bin/python",
          "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
          ↪ e-20260906/hormone-therapy/scratch/r03/host/attempt-01/world/pr
          ↪ ereg/execution/run.py",
          "--help"
        ],
        "model": null,
        "cwd": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referenc
          ↪ e-code-20260906/hormone-therapy/scratch/r03/host/attempt-01",
        "log": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referenc
          ↪ e-code-20260906/hormone-therapy/logs/r03-host-run-a01.log",
        "created_at": "2026-09-06T07:57:58.648961+00:00",
        "status": "finished",
        "pid": 2328413,
        "start_ticks": "81955356",
        "state": "R",
        "started_at": "2026-09-06T07:57:58.656055+00:00",
        "returncode": 0,
        "completed_at": "2026-09-06T07:57:59.276284+00:00"
      },
      {
        "label": "r03-host-run_tests",
        "command": [
          "/home/shaoerzhuo/anaconda3/bin/python",
          "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
          ↪ e-20260906/hormone-therapy/scratch/r03/host/attempt-01/world/pr
          ↪ ereg/execution/run_tests.py",
          "--output",
          "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
          ↪ e-20260906/hormone-therapy/scratch/r03/host/attempt-01/output"
        ],
        "model": null,
        "cwd": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referenc
          ↪ e-code-20260906/hormone-therapy/scratch/r03/host/attempt-01",
        "log": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referenc
          ↪ e-code-20260906/hormone-therapy/logs/r03-host-run_tests-a01.log",
        "created_at": "2026-09-06T07:57:59.286713+00:00",
        "status": "finished",
        "pid": 2328505,
        "start_ticks": "81955420",
        "state": "R",
        "started_at": "2026-09-06T07:57:59.293945+00:00",
        "returncode": 0,
        "completed_at": "2026-09-06T07:59:41.965310+00:00"
      }
    ],
    "completed_at": "2026-09-06T07:59:41.978105+00:00",
    "scope": "Structural coverage, syntax and synthetic execution; not
    ↪ scientific validation."
  }
}
```

```

},
"reviews": {
  "opus": {
    "status": "complete",
    "verdict": "accept",
    "report": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referen
    ↪ ce-code-20260906/hormone-therapy/reviews/r03/opus.yaml"
  },
  "codex": {
    "status": "complete",
    "verdict": "revise",
    "report": "/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-referen
    ↪ ce-code-20260906/hormone-therapy/reviews/r03/codex.yaml"
  }
},
"review_status": "revise",
"revision_slots_used": 2,
"continuation": "package_current_draft_with_unresolved_findings"
}

```

r01 / codex.md

Codex independent adversarial reference-code review — round 1

Verdict

Revise. The scientific implementation is substantial and its supported synthetic paths pass, but the execution/evidence trust boundary is not safe enough to call the reference-code snapshot faithful engineering. The required revision is narrow: authenticate predecessor acceptance, externally bind the executable code, and make manuscript slot selection enforce the same sealed aggregate envelope already enforced for figures. This is not a request for a new experiment or a change to the frozen protocol.

Scope and source-fidelity assessment

I reviewed the immutable original package, the read-only World, every protocol and result region for `atom-000001` through `atom-000005`, the complete Molecule, shared definitions and governance, the full D01–D20 deviation ledger, target catalog/inventory, source map/provenance, clinical-code and schema mappings, paper and 292 result-slot bindings, figure code, all execution code, the implementation report, and the complete supplied 969-line R source. The R file's SHA-256 was `a386ae0480f42705d5d8cf0b850f703123e962265e4d46715cecb2ecac0cfeea`, matching `source-provenance.yaml`. `diff -qr` showed that the reviewed World differs from the immutable original only by the added `prereg/execution` and `prereg/engineering` directories.

The implementation is not merely a wrapper that always reports missing data. The copied advertised suite ran 25 tests with no failures/errors, exercised all five Atoms as synthetic partial results, and invoked the permitted native Cox paths. Generated estimate identities exactly matched the target-catalog Cartesian products: 66

rows for Atom 1, 6 for Atom 2, 45 for Atom 3, 36 for Atom 4, and 27 for Atom 5. The 64 characteristic identities, 29 product identities, and all 292 unique manuscript placeholders also matched exactly, with no duplicate or missing binding.

I independently recomputed the primary crude model using a hand-coded Efron partial likelihood with the source counting-process risk rule `tstart < event_time <= tstop`. The result-code HR, confidence bounds, and P value matched within a maximum absolute difference of `4.0848e-7`. A first naive statsmodels PHReg run differed because the fixture has many event/change boundaries and that call did not reproduce the strict entry-at-event rule; I treated it as a limited oracle and replaced it with the explicit risk-set calculation, not as a code failure.

All five copied-layout governed CLI runs returned `coverage_state=boundary`, no operational terminal, and only `unresolved_source` estimate rows. This is correct. D01/D02 genuinely prevent an empirical cohort and risk set from being selected, so no real numerical result is currently available for any Atom. The lower-stage synthetic coverage does not resolve those source conflicts and is marked `synthetic=true`, `accepted=false`, and `registered_state=null`.

Findings

High — predecessor acceptance is forgeable by the manifest author

`prereg/execution/provenance.py:80-100` checks the bytes and internally self-consistent fields of a result, but the only acceptance credential is the result JSON's own `accepted` Boolean. The caller also chooses the absolute path in the manifest. There is no controller signature, controller-owned accepted-result lookup, or comparable host-authentication step, and `implementation_identity` is ignored. The author tests themselves construct such records at `prereg/execution/run_tests.py:417-427` and `451-457`.

My minimal falsifying record claimed `accepted=true`, `synthetic=false`, a registered boundary state, the correct material/input identity, and `implementation_identity=unapproved-reviewer-code`. I generated its digest and aggregate hash locally. `verify_evidence` accepted it. I then made the same kind of record claim an accepted `no_result` terminal for Atom 1 and passed it in a governed manifest. Even with `--dispatch-policy direct`, `provenance.py:104-112` returned activation false; the CLI emitted `terminal=not_run`, every estimate status was `not_run`, and no numerical field was populated.

The no-retry behavior is correct for authentic terminal evidence. The defect is that the evidence is not authentic. All five current activations are `TRUE`, but this still affects the current package because same-Atom terminal evidence is checked before the activation expression. It would also make any future evidence-dependent activation unsafe.

Narrow fix: resolve predecessor identifiers through a controller-owned accepted-results registry or verify a controller signature/token that is not inside caller-controlled result content. Bind the approved implementation identity as part of eligibility. A regression should reject an otherwise perfectly self-consistent unsigned

accepted/no_result fixture and show that direct dispatch remains active; it should retain the existing test that an authenticated terminal is never retried.

High — executable code is measured but not anchored or rejected

`prereg/execution/provenance.py:24-51` checks the World against `registration-lock.json`, which is located beside the executable. The lock's file list and `prereg/manifest.yaml:22+` contain frozen scientific materials, but no `prereg/execution` or `prereg/engineering` entries. `CHECKSUMS.sha256` likewise has no such entry. Lines 54-56 calculate the currently executing code's hash, and `run.py:76-80` records it in output, but no expected value is provided or compared.

In a copied World under reviewer scratch, I inserted one line at `models.py:48`: an immediate return from `target_gate`. The implementation identity changed from `aa65f872...` to `0d85b268...`, yet `verify_world` returned true. Running the copied Atom 5 CLI then crossed the frozen gates at `models.py:50-59`: the formerly prohibited adjusted `t3_other` nonreference cells were emitted as available (`0.554812`, `0.722374`), and Atom 5 reported `coverage_state=complete` rather than partial. The immutable reviewed World was never edited; this was entirely a copied-code counterexample.

Recording a changed implementation identity is useful provenance, but it is not rejection. A self-check inside the same mutable bundle cannot establish its own trustworthiness because the checker and co-located lock can be changed too. Narrow fix: have the controller/launcher validate the entire delivered execution bundle and lock against a digest held outside that bundle before importing `run.py`, or execute an immutable controller-addressed code object. Evidence eligibility must require the same approved digest. Add a one-byte estimator/source-gate mutation regression that must fail before outputs are created, while preserving the current legitimate tests for realized Atom body and Molecule interpretation changes.

Medium — result slots do not enforce the aggregate envelope

`prereg/execution/interfaces.py:8-43` directly opens the selected CSV/Markdown files and checks only selection shape and status presence. It does not open `result.json`, verify `result_digest` or `material_identity`, or compare the aggregate hashes. I copied the five passing Atom output directories and changed Atom 1's bound primary crude estimate from `0.87375280597817` to `999.0` without resealing the result. `select_slots` still returned all 292 slots and labelled `t2_primary_1_c` available. In contrast, `render_figures.py:14-24` correctly rejected the same Atom directory with `ValueError: Altered figure input`.

Narrow fix: factor the figure renderer's result/material/output validation into a shared result reader and call it for each Atom before slot selection. Add a regression where a selected aggregate byte is altered; selection must fail, while complete/partial/boundary/not-run sealed outputs remain selectable.

Per-Atom coverage

Atom	Reviewed implementation and executed coverage	Remaining numerical limit
atom-000001	Flow and exclusion union, end-of-follow-up versus age-55 summaries, Table 1, binary/duration/contraindications/models, calendar and surgery/BMI sensitivities, product and PH blocked interfaces. Synthetic result was partial; 66 estimate identities and 225 coverage rows were present. Reviewer strict-Efron oracle matched m01. Partial/missing/not-run figures were readable.	All empirical cells stop at D01/D02; adjusted publication attribution D08, PH D18, product D11, PH D13, and applicable D03–D07/D09 remain unavailable.
atom-000002	Full-universe parent cleaning, maternal eventual-discordance siblings, retained histories, independent population/count/PY/rate rows, and stratified native-math test. Synthetic result was partial; 6 estimate identities and 8 coverage rows were present.	Empirical cohort D01/D02; printed crude within-family attribution D11 and adjusted attribution D08 remain unavailable. No clustered sandwich substitute was introduced.
atom-000003	Hysterectomy and BSO45–54 postoperative risk sets, cancer censoring, crude models, BSO death count and conditional death-age quantiles. Synthetic result was partial; 45 estimate identities and 52 coverage rows were present.	Empirical cohort D01/D02/D05; older-BSO and code hysterectomy-age targets D12; adjusted publication cells D08 remain unavailable.

Atom	Reviewed implementation and executed coverage	Remaining numerical limit
atom-000004	Persistent exposure, duration, strict route/regimen dominance, tie failures, and independent initiation branch. Synthetic result was partial; all 36 estimate identities/coverage rows were present.	Empirical cohort and episode units/construction D01–D04; route/regimen ties D10; adjusted attribution D08 remain unavailable.
atom-000005	Separate cause horizon interface, known-cause classification, competing deaths censored, crude cause-specific Cox fits, summaries, and common-denominator cause proportions. Synthetic result was partial; 27 estimate identities and 30 coverage rows were present.	Empirical cohort/horizon/cause version D01/D02/D09/D15; adjusted attribution D08 and absolute other-adjusted D14 stop remain unavailable. This Atom supplied the code-tampering counterexample above.

The R estimator matches the source’s counting-process `coxph`, Efron ties, and model-based covariance. There is no RD, staggered DiD, matching, cluster-robust estimator, frailty, or competing-risk cumulative-incidence model in the frozen source, so their absence is not a fidelity defect. Other classified deaths are treated as censoring for cause-specific hazards, as registered.

Executed tests and limits

The reviewer-executed author suite command was:

```
/home/shaoerzhuo/anaconda3/bin/python -B
↪ world-copy/prereg/execution/run_tests.py \
  --output /home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-cod
↪ e-20260906/hormone-therapy/scratch/r01/codex/baseline-suite.aHuBjT
```

It ran 25 tests in 69.654 seconds with zero failures/errors, reported the World unchanged, and explicitly reported `pinned_engine_validated=false`. An earlier invocation without `--output` exited 2 at `argparse`; that was a reviewer usage error. Reviewer-owned checks are replayable with:

```
/home/shaoerzhuo/anaconda3/bin/python -B \  
/home/shaoerzhuo/Bolero-denmark/.prereg-runs/denmark-reference-code-2026090_\  
→ 6/hormone-therapy/scratch/r01/codex/adversarial_checks.py
```

That command records the strict Efron agreement and all three counterexamples. The exact catalog/binding audit is replayable with `catalog_audit.py` in the same directory. Supporting copied outputs, mutated code, forged evidence, governed CLI artifacts, and visual-check rasters remain only in reviewer scratch.

Frozen R 4.4.1/survival 3.7-0/data.table 1.16.2 were unavailable. Installed R 4.1.2/survival 3.8.3 executed the synthetic native checks; the implementation correctly refuses a non-synthetic engine mismatch. `survminer` and the exact `ggcoxzph` transform/smoothers were unavailable, so no substitute PH numerical output was accepted. No confidential data, old empirical session, extraction equivalence, disclosure decision, QGPU, empirical result, or publication Helvetica Neue bundle was tested. These limitations leave all real study numerical results unavailable; they do not excuse the three engineering defects or turn source absence into a scientific null.

r01 / opus.md

Independent engineering review — Opus, round 1

Verdict: revise (four narrow, located code fixes plus regression tests). The verdict rests on **F2** alone — a confirmed numerical divergence from the source’s own R136 expression in a named kernel — and does not depend on the other three findings.

Scope: the frozen reference implementation under `v001/world/prereg/execution/` and `v001/world/prereg/paper/figures/`, audited against the five frozen Atom protocols, the Molecule, `prereg/data-contract.md`, `prereg/replication/*` and the 969-line source R file (`mikkelsen-supplement2.txt`, sha256 `a386ae04...c0cfeea`, hash re-verified). This is an engineering/source-fidelity review only. It is not scientific acceptance of the frozen draft, not approval of the original protocol, and not a replication: **no empirical numerical result of any kind exists or was produced**.

Everything I ran was written under `scratch/r01/opus/` (a writable copy of the World in `copy/world`, the author suite output in `tests-out/`, my own tests in `mytests/`, their output in `myout/`, PDF previews in `previews/`). The reviewed World is byte-identical: `sha256sum -c prereg/CHECKSUMS.sha256` passed with no mismatch and no file under `v001/` was modified. I did not read the other reviewer’s outputs (`reviews/r01/` contained only my own file), did not spawn subagents, did not touch real data, and made no commit/push/release.

1. Is this reference code or a wrapper?

Reference code. `construction.py` (432 lines) implements the source’s actual kernels; `engine.R` makes real `survival::tmerge / survSplit / coxph` calls; `analysis.py` exports the exact frozen component roster; `models.py` holds the literal source formula table. Nothing is a TODO.

The one structural caveat is honestly declared and, on my checking, genuine: `Analysis.shared_gate` (`analysis.py:51-68`) stops **every** target with `STOP_UNRESOLVED_SOURCE` D01/D02 whenever `purpose != "synthetic"`, and `run.py:105` never even loads frames in that mode. So the governed path can produce no number at all. I confirmed D01 is a real blocker rather than an access excuse: source line 75 is `start:=as.IDate(pmax(fdato+(45*aar), as.IDate(1995-01-01), first_ind, na.rm=T))` where `1995-01-01` is **unquoted** — numeric arithmetic evaluating to 1993, i.e. 1975-06-16 — so the source entry rule cannot be executed as written, and D02’s migration rule (R51-53 vs P3’s “six months”) is likewise underdetermined. Every target in this study is cohort-derived, so stopping all of them is the faithful consequence, not an over-broad gate.

Crucially, the supported analysis mathematics is still implemented and executed on explicitly synthetic logical inputs, and **target-level** conflicts stay closed even there (`models.target_gate`, `models.py:48-59`). My own run produced these coverage rows, with **zero operational failure**:

Atom	rows available		reference	unresolved_source
atom-000001	225	154	12	59
atom-000002	8	4	0	4
atom-000003	52	17	2	33
atom-000004	36	21	3	12
atom-000005	30	18	3	9

Untouched executable siblings are retained exactly where the ledger permits: the code-only adjusted fits `m112/m122` (R640/R654), `m132` (R663) and `m162` (R705) run, because D08 concerns the *published* Table 2/3 attribution only; `code_smoking`’s summary runs (R714-715) while its D13 fits stop; `t2_hyst` and `t2_bso_45_54` survive while D12 closes the `>=55` blocks.

2. Per-Atom coverage and what I checked

atom-000001 — Population mortality

Implemented and numerically exercised: `t2_primary` (m01, R458), `t2_duration` (m11, R474), `t2_contraindication` (m21, R493 with the R486 row filter), `t2_cessation` (m41, R535), `code_calendar_pre2003/post2003` (m111/m112, m121/m122, R637-R656), `code_surgery_adjusted` (m132 with the shared `rt01` crude, R663/R667), `code_bmi` (m161/m162, R702-R707), `code_smoking` summary only (R714-715), the 17-node `flow.csv`, all 32 Table 1 cells, and every descriptive scalar.

I verified each descriptor against its exact source line rather than the protocol prose: follow-up quartiles = per-woman `sum(risktime)` quantiles (R817, type-7); `uptake/cohort_deaths`/birth-cohort shares from the last row per woman (R806, R819-R827); `end_age` = median final `tstop` (R454/limitations, a descriptive quantity, not life expectancy); the landmark uses `tstart==55` (R950) while flow and miss-ingness use the last `tstop` (R963-R967 map exactly onto `uddniv2=="4_Unknown"`,

`civst1=="3_Unknown",fland1=="3_Unknown",q_ind=="q5")`; landmark uptake windows [2004,2007) and [2021,2024) (R957-960) rather than an invented trend; the contraindication-death share reproduces R489. Flow counting matches R280-R299: eligibility counted before exclusions, the seven overlapping reasons counted individually, and the union counted separately; `.le(start_age)` reproduces `data.table`'s NA-dropping `<=start_age` filter.

Stopped and correctly so: m02/m12/m22/m42 (D08), the 14 PH panels (D08/D18 — `blocked_displays` emits all 14 with reasons and empty residual/curve/tick tables), the 29 eTable 2 rows (D19 — I confirmed R430 sums `risktime` on rows with non-missing `htatc`, which is not S2's "proportion of MHT"), smoking fits (D13 — R717/720 fit `cpr8bmi` and R725 merges `rt161`).

atom-000002 — Sibling mortality

`construction.siblings` reproduces R503-R507 precisely: parent >30 cleaning on the **full** person universe (R45-47) before cohort restriction, sisters counted within the eligible cohort, maternal groups retained on eventual `max(ht_sib)=1 & max(1-ht_sib)=1`, no shared-father requirement, then an inner many-to-one join back to every original interval. My test confirmed no row multiplication and that a later-exposed sister still contributes her unexposed time. Crude m31 stops at D11 (R515 has no `strata`, P3 says mother strata) and no strata were added; adjusted m32 keeps the exact R518 formula (including `strata(family_id)`) but stops at D08. The README correctly states the source has model-based, not family-clustered, covariance — I verified this numerically (§4).

atom-000003 — Surgical subgroup mortality

`t2_hyst = breastgyn==0 & hyst==1` with collapsed cuts (R569); `t2_bso_45_54 = breastgyn==0 & bso==1 & bso_age<55` (R583-R587). I tested the boundary directly: a woman with `bso_age == 55.0` is excluded from the younger group (she belongs to R583's `bso:=2` overwrite), and preoperative intervals never enter the risk set. `bso_deaths/bso_death_age` reproduce R597-R599 as *conditional-on-dying* `tstop` quartiles, correctly separated from the population N. `t2_bso_ge55`, `code_hyst_1t55` and `code_hyst_ge55` stop at D12 for the retrospective-grouping conflict, and the younger results survive independently — which is the granularity behaviour the Atom rationale promises.

atom-000004 — Treatment heterogeneity

`exposure_intervals` reproduces R387-R427: `ddd` join on (`pnr`, `htstart==tstart`), missing dose → 0, `ht_sum = cumsum(ddd)/365.2`, persistent ever-use, strict cumulative maxima for route and regimen, `estrogen+progestin` split into monophasic/cyclic on Jaydess|Kyleena|Mirena|Levosert (R410). I confirmed a tie leaves the category unassigned and does **not** carry across persons (R401/R422's `nafill` without `by=pnr` is deliberately not reproduced — that is D10) and that `prepare_target` then raises D10 rather than guessing. `initiation` (`construction.py:390-395`) reproduces the **retained**

R618 recipe, not the overwritten R616: pre-exposure rows fall in the never-use category, and I verified the $[0, .1) / [.1, 52) / [52, 57) / [57, 150)$ boundaries at exactly 52 and 57. This is the one place I found a source-fidelity defect in the route branch (F3).

atom-000005 — Cause-specific mortality

`cause_classification` refuses to resolve R745's cancer-then-cardiovascular overwrite precedence or the `is.na(csm) -> "other"` rule (D15), and I confirmed it does **not** silently add cause code 12 to cancer. `cause_intervals` sets `event = event & cause==target`, so other classified deaths censor at their own death time without any risk time being subtracted — I checked the person-years explicitly. No Fine–Gray, cumulative incidence or all-cause subtraction exists anywhere. `cause_proportions` reproduces R797-R799's three-cause denominator. Adjusted cardiovascular m202 keeps R756's reduced covariate set `A_CV` but stops at D08; `t3_other` adjusted stops absolutely at D14 (R785 extracts m92), and I confirmed neither m92 copying nor m222 substitution is reachable.

3. Cohorts, clocks, leakage, ties, joins, units

- **Index selection vs time zero vs later events** are kept distinct: landmark `tstart==55`, flow/uptake on last `tstop`, follow-up as summed interval length, death age as final `tstop`. The README states this explicitly and the code matches.
- **Immortal time / future information**: the source's own designs (eventual sibling discordance, backward-constant `first_ht` category, cumulative route) are reproduced without adding any further look-ahead, and `prepare_target` (models.py:68) rejects an ever-use indicator that returns to never. `initiation` assigns 0 (never) to pre-exposure rows exactly as R618 does.
- **Ties**: route/regimen ties stay unresolved (D10); income-quantile duplicate breaks and ambiguous `alder_ult_ink` ties stop rather than picking a rule (D07); Efron tie handling is the source's, verified numerically.
- **Duplicate joins**: `join_many_to_one` enforces right-side uniqueness with `validate="many_to_one"`; `validate_intervals` rejects duplicate (`pnr, tstart`), overlapping histories, multiple deaths and post-death rows.
- **Units**: `estimates.csv` labels hazard ratio vs events/person-years/deaths-per-10000-person-years; years everywhere; formatting is confined to `interfaces.present_estimates` and never feeds the machine table.
- **Longitudinal dependencies** are where I found F1 and F2 (below).

4. Estimator and inferential fidelity

`engine.R:44-51` calls `coxph(..., ties="efron", robust=FALSE, na.action=na.fail, conf.int=.95)` — the pinned survival 3.7-0 defaults for calls that omit those arguments. There is no RD, DiD, matching or alternative survival model in this source, and none was substituted.

My independent oracles:

- I wrote my own Efron partial likelihood and optimised it with SciPy for a three-level `ht_cat2` crude fit; the engine’s HRs agreed to 5 decimals.
- `statsmodels PHReg` reproduced the model-based SE on 40 women $\times$ 2 intervals to 6 decimals, confirming **no** clustered sandwich is silently applied — matching the source, which has none.
- I verified in R that a `strata()` term changes baseline hazards only.
- Named coefficient extraction (`models.py:179-183`) replaces R’s fragile positional `conf.int[1:5,]` recycling (R475/R551), which is meaning-preserving and is the right response to D16; failed coefficients stay blank and reference rows carry no fabricated CI.
- Source–author code disagreements are preserved, not repaired: D08, D11, D12, D13, D14, D16 and D19 all reproduce exactly as the ledger describes when read against the R file.

5. Findings

Full evidence, fixes and regression tests are in `opus.yaml1`. In short:

ID	Severity	Where	What
F1	moderate	engine.R:17–21	The change-variable whitelist omits R319’s cancer tmerge. Supplying it $\rightarrow$ RuntimeError $\rightarrow$ whole-Atom no_result instead of a named boundary; omitting it moves person-time across the R633/R647 year_start<2003 filter, which D09 records has no aligned split point (5.0/10.0 $\rightarrow$ 3.6/11.4 person-years in R).

ID	Severity	Where	What
F2	major	<code>construction.py:139-141</code>	R136 on picks the terminal episode by last <code>htstart</code>, but R136 uses <code>order(pnr,-htstop)</code>. On nested episodes — reachable through the package’s own <code>prescriptions</code> → <code>episodes_unambiguous</code> chain, which splits on product name — I measured <code>htstop90</code> = 51.831 versus R’s 56.182. This shifts the <code>t2_cessation</code> risk set.
F3	minor	<code>construction.py:385-386</code>	R136 form outside R94-R100’s map is turned into a missing route, whereas R394 deterministically classifies it as “other formulation” (confirmed in <code>data.table: ht_dosform 3</code>). The stop is reported under D10, which covers tie carry, not unmapped forms.

ID	Severity	Where	What
F4	minor	<code>models.py:153-154</code>	A single-observed-level adjustment factor stops the fit, though R fits it with only the absent level aliased (verified: identical <code>ht</code> coefficient) and the code's own <code>absent_terms</code> logic at <code>models.py:170-177</code> already covers this case.

A claim I withdrew. My first F1 fixture supplied only `dm` and `cancer` changes and appeared to show 3.6 person-years moving between `year_cat` categories. That was a fixture artifact: R302-305/R322-325 tmerge `age2002...age2023`, and engine.R *does* whitelist all four, so those splits land on the `year_cat` boundaries. Re-running with them present, person-time per `year_cat` is identical with and without the cancer `tdc` (4/7/6 both ways). I withdraw the `year_cat` claim; `m132/m162` are not first-order affected, and F1 is scored moderate rather than major. What survives is the interface defect plus the R633/R647 filter effect, which is a direct consequence of D09's recorded "no 2003 split". A small second-order `year_cat` effect remains possible because ages use the 365.2-day clock while `year_start` uses `decimal_date`, but I do not rest the finding on it.

F2 changes a scientific quantity in a currently-implemented target, which is why the verdict is `revise` rather than `accept`. F1 additionally converts a registered construction input into an operational failure. F3 and F4 are conservative in direction (a stop, never a fabricated number) and could alternatively be resolved by documenting them as explicitly converter-added restrictions.

Two adjacent things I checked and decided are *not* findings

The extra `n` column. `Analysis.models` adds `n = y.pnr.nunique()` to summary rows, which R456 does not compute (it produces only `event`, `risktime`, `ir`). I checked where it can surface: of the 180 estimates slots, 120 select `estimate,lower,upper,p_value,model_id` and 60 select `events,person_years,rate_per_10000` — `none` selects `n` — and `published-reference-values.csv` holds no women-count comparator for those targets. It is an unbound internal audit field and cannot reach the manuscript or a published comparison.

<3-event suppression. R729 does `[event<3, event:=NA]` — `event` only — while keeping `risktime` and `ir`, from which `events = rate × person_years / 10000` is exactly recoverable. The code withholds `n`, `events`, `person_years` and `rate_per_10000` together (`analysis.py:120-122`). That is *more* than the source does, so it belongs to the same class as F3 and I owe a reason for treating it differently: unlike F3 it is (i) written into the atom-000001 protocol step 11 as source disclosure treatment, (ii) declared verbatim in `execution/README.md`, and (iii) strictly necessary for the registered rule to have any effect at all. It is disclosure engineering implementing a registered rule, not an unregistered scientific restriction filed under someone else’s deviation ID. F3, by contrast, withholds a source-*determined* classification with a published comparator and reports it as D10, which covers tie carry only.

6. What I confirmed is faithful

Dispatch and evidence semantics are correct. With `purpose: governed, --dispatch-policy direct, definitions: {D01: resolved, D08: "adjustment set = A_P", resolved: true}`, a hashed `source_evidence` item and `disclosure: true`, every component still returned `unresolved_source` with all numeric fields blank. No fake prior result is accepted (`verify_evidence` rejects synthetic or unaccepted evidence, wrong material/input binding and corrupt aggregate bytes); an explicit `no_result` terminal is never retried under either policy; contradictory branches cannot be forced; unknown activation yields `not_run` rather than execution. The native parser/evaluator is byte-identical to the repository’s `prereg_lane/native.py` (I diffed function by function), so `NOT unknown` stays unknown and `done()` requires a realized state.

The Molecule policy holds and is interface-compatible with the executor. Membership covers all five Atoms (atom-000001 primary, the other four secondary), each with its own scientific rationale. Every interpretation condition references only local verdict keys that actually exist in the member Atoms — `boundary`, `complete`, `partial` — plus `done()`; a blanket `TRUE` restriction is evaluated first, per-Atom boundary restrictions precede the primary/supplementary permissions, and a `TRUE` residual reports unresolved scope. The two joint-reading rules explicitly forbid a difference test, a pooled effect or an equivalence rule, and forbid one available result standing in for the unresolved older-BSO or other-cause adjusted targets. This matters for the code because `analysis.run` (`analysis.py:285-291`) can only emit `boundary`, `partial`, `complete`, `otherwise` or `None`, all of which those rules or the residual can consume, and `run.py` never writes `registered_state` — host adjudication is preserved.

Registration identity survives legitimate realization changes and rejects tampering: appending to an Atom body and to the Molecule’s current interpretation left `material_identity` unchanged and altered `world_realization`, while editing a registered verdict region was rejected. The copied action directory ran against a `chmod 0555` relocated World with output outside it; `manifest_identity` ignores physical dataset paths and rejects absolute paths inside scientific metadata.

Missing inputs are never disguised as science: a missing `Rscript` yields `boundary`, a failing one yields terminal `no_result`, available sibling summaries cannot upgrade a `no_result` artifact to `partial`, and a group with 2 events is

suppressed with events/person-years/rate all blank so the count is not recoverable. `source-comparison.csv` stays an empty declared interface and no published value is ever copied into an output.

Manuscript and figure interfaces hold: all 292 fixed slots select against generated tables in both the complete and the fully unactivated dispositions; `paper.md` uses exactly those 292 slot names with none undefined or unused and contains no fabricated HR/CI text. I rendered and visually inspected the complete, partial and all-missing figure PDFs at 130.7 mm: figure 1's long unavailable-reason note is complete and unclipped and its flow sums are internally consistent; figure 2 keeps all 14 source-ordered terms with nine numbered explanations, no overlap or clipping. Both templates raise rather than render when a value is attached to an unavailable status.

7. Numerical results that remain unavailable, and why

Nothing in Table 1, Table 2, Table 3, Figure 1 or supplementary Figure 1 is currently obtainable. Under **purpose: governed** the code stops before reading any person row because D01 (R75's unquoted `as.IDate(1995-01-01)`, plus the 365.2-day versus calendar-birthday clock) and D02 (R51-53 migration ordering and the 180-day versus six-month rule) leave the cohort and follow-up undefined. Beyond that shared block: every published adjusted Table 2/3 estimate is closed by D08 (R461 and all adjusted calls include `civst_mode` while the abstract and the Table 2 footnote omit civil status); `t2_bso_ge55` and both `code_hyst` age blocks by D12; `t3_other` adjusted absolutely by D14; `code_smoking` crude and adjusted by D13; the 29 eTable 2 proportions by D19; the 14 Schoenfeld panels by D08/D18; route and regimen by D10 whenever cumulative maxima tie. These are faithful blocked targets, not failures of this implementation.

Software limits: the pinned R 4.4.1 / survival 3.7-0 / data.table 1.16.2 environment is **not installed here** (I ran R 4.1.2 / survival 3.8.3 / data.table 1.18.2.1), `survminer/ggcoxzph` is unavailable so no Schoenfeld smoother could be validated, and publication Helvetica Neue resources are absent so figure checks used the documented DejaVu Sans fallback. `Engine(..., synthetic=False)` correctly refuses to run on the unpinned engine, so no non-synthetic fit was executed by me or by the author suite.

8. Tests I ran

- **Author suite, re-executed by me** against a copy of the World: `run_tests.py --output <scratch>/tests-out` → 25 tests, 0 failures, 0 errors, 71.9 s, `world_unchanged: true`, `pinned_engine_validated: false`. This matches the author's final-validation record.
- **My own 15 adversarial tests** (`mytests/reviewer_tests.py`): 12 passed, covering initiation/duration cut boundaries at 52, 57, 5 and 100; route ties and no cross-person carry; the BSO preoperative and age-55 boundary; a hand Efron partial-likelihood oracle for a 3-level factor; a non-cluster-robust SE oracle; meta-data that cannot clear a frozen stop; the full target-gate matrix including the

code-only siblings that must stay open; <3-event suppression; competing-cause censoring without risk-time loss; unclassified/conflicting causes and cause code 12; sibling history preservation; and realized-body mutation versus verdict tampering. The 3 failures are findings F1–F3.

- **R oracles** executed against the frozen source expressions: the R319 tmerge comparison on the R633 calendar filter, its `year_cat negative control` (which falsified my own first version of F1), the literal R136 cessation expression in `data.table` (both the minimal and the pipeline-generated case), the R390-R399 dose-form assignment, and a constant-factor `coxph` for F4.
- **Static check:** AST function-by-function diff of `prereg/execution/native.py` against `tools/prereg-lane/src/prereg_lane/native.py` — identical.
- **Visual inspection:** `pdftoppm` renders of the six synthetic figure PDFs.
- **Integrity:** `sha256sum -c prereg/CHECKSUMS.sha256` clean; no file under `v001/` modified.

9. What revise asks for

Four narrow code fixes and their regression tests, listed with exact locations in `opus.yaml`. No new research question, cohort, estimator, threshold, robustness check or Atom is requested, and no change to the frozen protocol, roster, deviation ledger, result slots or manuscript is required. With F1 and F2 fixed (and F3/F4 either fixed or explicitly re-documented as converter-added restrictions), I would expect this package to qualify as faithful, tested engineering within its explicitly stated and, in my checking, genuine limitation that no empirical numerical result is presently obtainable at all.

r02 / codex.md

Codex round-two independent engineering review

Verdict: **revise**.

This snapshot is a genuine reference implementation, not a wrapper that merely reports absent data. From a copied layout I ran the complete advertised suite: 32/32 tests passed with no errors and no World changes. Its synthetic outputs covered all five Atoms, 351 component rows, 20 structural references, 214 available rows, 117 explicit source stops, 20 native model invocations, and all 292 fixed manuscript slots. I separately executed every governed Atom with the empty registered manifest. Each produced a boundary artifact, every numerical estimate field was blank, and all slots remained selectable as missing dispositions. That is faithful blocking, not a null or equivalence result.

The round-one scientific-code repairs held in my targeted cases. R319 cancer is present in both tmerge whitelists and the literal hospital/cancer-file predicates exclude C44/191 while retaining other C/numeric cancer codes ([construction.py:157-177](#); [engine.R:17-30](#)). The nested cessation case returned 51.246 from an earliest qualifying gap despite a later-starting episode, matching literal R136 ordering ([construction.py:125-139](#)). An unmapped nonmissing SUSP form contributed to route

`other`, while a subsequent equal cumulative other/transdermal total remained unknown, as required by R394 and D10 ([construction.py:364-396](#)). Constant/absent adjustment-factor aliases were exercised by the supplied suite.

The Codex round-one acceptance and code-pin repairs also mostly held. A caller-only result ID could not control dispatch; an externally registered exact result byte hash is required at [provenance.py:99](#). Strong Kleene NOT `unknown` remained unknown. An edited `models.py` in a copied action bundle was rejected by the standalone launcher before output, and the recomputed bundle/launcher hashes matched the author report. The incomplete fix is the separate manuscript/figure consumer described below.

Findings

C2-01 — High: canonical consumers accept caller-resealed results

`read_result` validates only public, recomputable hashes and a result-supplied implementation ID that is a member of the approved-code list ([provenance.py:76](#)). It does not bind the exact `result.json` bytes to the external controller registry. Both the slot selector ([interfaces.py:9](#)) and figure adapter ([render_figures.py:13](#)) use this weaker reader.

I began with actual outputs generated by this snapshot, changed the atom-000001 primary crude HR to 999 and its eligible flow count to 999, recomputed each CSV SHA256 and `result_digest`, and left the genuine approved implementation digest in the envelope. The selector accepted all 292 slots and the renderer produced a PDF containing the forged flow input. The same result could not act as predecessor evidence, because that path correctly requires the controller-bound result byte hash. This precisely isolates the remaining gap.

The narrow fix is to require an externally authenticated exact-result binding (or equivalent protected attestation) for canonical slot and figure consumption. Synthetic/unaccepted previews can remain available through an explicitly noncanonical preview path. A regression must rehash and reseal an edited generated aggregate and confirm canonical consumers still reject it.

C2-02 — Medium: a declared governed dataset need not have a content hash

`load_manifest` checks path/provenance/schema fields but not `sha256` ([run.py:26](#)). Hash verification occurs only in `load_frames`, and governed execution deliberately never calls that loader before D01/D02 blocking ([run.py:99](#)). Consequently a declared `intervals` dataset with no hash and a nonexistent path was accepted and received an input identity and boundary result. This contradicts the documented dataset descriptor and “corrupt hashes or malformed manifests reject” contract.

Require a 64-hex content hash for every declared dataset during manifest validation, without opening person rows. An actually absent optional input should be omitted from `datasets`. The empty governed manifest must continue to emit its legitimate frozen boundary.

C2-03 — Medium: lower-stage source domains and longitudinal premises are under-validated

Only `ht` receives an explicit binary-domain check. The executable `m132` adjustment path accepts source `tdc` indicators such as `dm` as unrestricted numeric values ([models.py:11](#) and [models.py:146](#)). Changing selected `dm` values to 2 still produced a supported `code_surgery_adjusted` coefficient. That is outside the source logical interface, where these variables arise as `tdc` indicators.

Separately, `validate_intervals` permits unexplained gaps ([construction.py:339](#)), while `source_calendar` advances by cumulative risk time and therefore assumes the upstream `tmerge` history is contiguous ([construction.py:304](#)). For (45,46] and (50,51] with calendar entry in 2000, it silently assigned the second row to 2001 rather than the interval clock's 2005. The literal cumulative formula is correct for the source's lossless contiguous splits; the accepted interface fails to enforce that prerequisite.

Add component-scoped `{0,1}` checks for consumed disease/surgery indicators and reject unexplained gaps before calendar kernels that rely on contiguous source histories. These checks should not discard independent outputs that do not consume the invalid optional field.

Per-Atom coverage

- `atom-000001`: 86 registered targets, 170 slots. Synthetic execution produced 225 coverage rows: 154 available, 12 reference, 59 unresolved-source; 12 native model audits. Governed execution produced 225 D01/D02 boundary rows. The `m01` tied-Efron coefficient and model-based SE matched my independent risk-set calculation: beta -0.027500497567 versus -0.027500497250 , SE 0.244317230680 versus 0.244317230686 . C2-01 affects its slots/figures; C2-02 and C2-03 affect its runtime contract.
- `atom-000002`: 2 targets, 8 slots. Synthetic execution produced 8 coverage rows (4 available, 4 unresolved-source); the D11 crude and D08 adjusted fits remained stopped while sibling summaries ran. Governed execution produced 8 boundary rows. C2-01 affects its slots and C2-02 is shared.
- `atom-000003`: 7 targets, 48 slots. Synthetic execution produced 52 coverage rows (17 available, 2 reference, 33 unresolved-source) and 2 native crude model audits. Younger surgical outputs survived the D12 older-surgery stops. Governed execution produced 52 boundary rows. C2-01 affects its slots; source-binary validation in C2-03 also covers its surgery indicators.
- `atom-000004`: 3 targets, 37 slots. Synthetic execution produced 36 coverage rows (21 available, 3 reference, 12 unresolved-source) and 3 native crude model audits. Route-other/tie and initiation remained independent. Governed execution produced 36 boundary rows. C2-01 affects its slots and C2-02 is shared.
- `atom-000005`: 4 targets, 29 slots. Synthetic execution produced 30 coverage rows (18 available, 3 reference, 9 unresolved-source) and 3 native crude model audits. My competing-risk case retained person-time and changed the competing death to event zero. Governed execution produced 30 D01/D02/D09/D15 boundary rows. C2-01 affects its slots and C2-02 is shared.

Limits and executed evidence

All empirical numerical outputs remain unavailable. D01/D02 block the cohort and risk sets; D03-D07 block unverified source mappings and construction; D08-D15, D18, and D19 continue to stop only their consuming adjusted, horizon, surgical, smoking, cause, diagnostic, and product components. I found no basis to narrow or override those registered stops.

The controller’s frozen R 4.4.1/survival 3.7-0/data.table 1.16.2 environment was unavailable. Native synthetic checks used installed R 4.1.2/survival 3.8.3, and non-synthetic inference remains correctly version-gated. I did not access real data, old empirical sessions, disclosure state, QGPU, or controller deployment. survminer/source-compatible ggcoxzph and publication Helvetica Neue resources were unavailable; no substitutes were accepted.

The full reviewer-owned evidence is in [independent-report.json](#) and the supplied-suite rerun in [test-report.json](#). The adversarial harness itself is [independent_checks.py](#). The harness reports 16 passing checks and five intentionally failed rejection assertions; its exit 0 records completion of the evidence report, not acceptance of those five behaviors.

r02 / opus.md

Round-two independent engineering review — Opus

Verdict: accept (with two minor located code fixes, F5 and F6)

Reviewed World: v002/world. Exclusive scratch: `scratch/r02/opus`. This is an engineering / source-fidelity review. It is not scientific acceptance of the frozen protocol, not approval of the original study, and not a replication. No new experiment is requested anywhere below.

0. What actually changed in v002, and what did not

`diff -rq` between the immutable package (`.controller/source`) and v002/world reports only two added directories, `prereg/execution` and `prereg/engineering`. All 78 entries of `prereg/CHECKSUMS.sha256` verify. `diff -rq v001/world v002/world` shows changes confined to `prereg/engineering/implementation-report.md` and nine files under `prereg/execution/`, plus the new `launcher.py`.

Nothing in `atoms/`, `molecules/`, `prereg/replication/`, `prereg/paper/` or `prereg/diagrams/` moved a byte. So the Molecule policy, the activation truth tables, the Atom roster, the deviation ledger, the target catalogue, the 292 result slots and the two figure templates are the same objects I read in full in round one, and my round-one confirmations about them carry forward unchanged. I re-read the changed code in full and re-derived every round-one finding with my own instruments rather than trusting the author’s tests.

I ran the advertised suite first: **32 tests, 0 failures, 0 errors, world_unchanged: true** (86 s). I also ran it *without* the documented `-B`, because `launcher.bundle_identity` rejects any `.pyc` in the bundle — `run_tests.py` sets

`sys.dont_write_bytecode` at line 4, so no `__pycache__` is written into the World and there is no self-brick. No defect.

1. Round-one findings: each one re-tested independently

Opus F1 — omitted R319 cancer tmerge (was moderate)

My round-one fix asked for one word in a whitelist plus documentation. The author added that **and** a new clinical construction, so I audited the construction against the frozen source (`mikkelsen-supplement2.txt`, sha256 `a386ae04...`, 969 lines) rather than assuming.

- `engine.R:18-21` and the new Python pre-check `models.py:133-138` are **set-identical to each other** and **exactly equal** to the 24 tmerge variables the source declares at R309–R332 — no omission, no extra (`mytests/t_static.py`). The Python check also converts what used to be an R `RuntimeError` → `operational_failure` → whole-Atom `no_result` into a named Boundary(`"unavailable_input"`, `"Unregistered change variable"`).
- `cancer` appears in **no** SPECS adjustment set and is **not** in `FACTOR_LEVELS`. The full adjustment-term inventory is unchanged: `year_cat`, `fland_mode`, `civst_mode`, `udda_mode`, `q_ind_mode`, `parity_cat`, `dm`, `hypertens`, `afli`, `valv`, `hf`, `statin`, `many_diag`, `bmi_cat`, `ryger`, `hyst`, `bso`, `strata(family_id)`. So it is a split point only, as my fix required.
- `construction.py:165-171` uses only verbatim source predicates: R198 `~DC[0-9]` minus `~DC44` on the hospital frame, R218 `~C|^ [0-9]` minus `~C44|^191[0-9]` on the cancer-register `search` field, unioned and reduced to the per-person minimum date per R219. I built a **literal R198/R213/R218/R219 data.table oracle** and ran it over 50 synthetic persons and 29 adversarial code strings (`C44`, `C441`, `C50`, `191`, `1910`, `1919`, `191A`, `1900`, `c50`, `DC44`, `DC441`, `DC0`, `DC97`, `DZ00`, ...): **0 mismatches**.
- Missing historical cancer still yields the D05 Boundary, the same treatment as `breast_gynaecological` — correct, because R219 needs both sources.
- Documentation: `execution/README.md`'s `changes` provenance row states cancer is “retained solely as a split point ... no cancer covariate or new outcome model.” That is the D20 / `code_unused_outcomes` crosswalk my round-one fix asked for.

One note I record but do not raise as a finding: this is the only clinical predicate in the codebase written as a literal instead of read from `clinical-codes.yaml`. That file is frozen and checksummed and has no `diagnoses.cancer` entry, so the author could not have put it there without tampering with registered material. Given the regexes are verbatim and documented, the placement is the correct choice under the constraint.

Opus F2 — cessation terminal episode (was major)

`construction.cessation:130-139` now marks `x.htstop.eq(groupby("pnr").htstop.max())` as the fallback. I checked the two things the author's own oracle test structurally cannot check, because it feeds identical CSV bytes to both engines:

- The function sorts by (`pnr`, `htstart`) on entry and enforces uniqueness on that key, so the `shift(-1)` lead is order-deterministic. That matches R136’s ordering of `ht4`, which R125 (`order(pnr, min_eksd)`) and R132 (`.SD[1]`, `split`) leave in (`pnr`, `htstart`) order.
- Hand case with the maximum-end episode in the **middle** of a shuffled three-row input: matches. **150-person fuzz** with 1–5 episodes each, nested/overlapping/tied ends and spans at exactly $90/365.2 \pm 1e-6$, rows deliberately shuffled: **0 mismatches** against the literal R136 expression.
- Tie reasoning holds: R’s `n1==1` marks exactly one row while the Python mask marks all tied maxima, but equal `htstop` equal `htstop90`, so the selected value is identical.

Opus F3 — unmapped dose form (was minor)

The added `unknown_form` nullification is deleted. I confirmed from the source that R394 operates on the **mapped** `dosform` (R94–R100 assign lowercase `oral/transderm/...`), so an unmapped string legitimately falls into the strict-maximum “other” arm. Oracle over 160 interval rows, 40 persons and 12 dose forms including `SUSP`, `PULV`, `SOURCE_UNMAPPED`, `TABORAL`, `transdermNEW`: **0 mismatches** against literal R390–R405 `ht_dosform`, including every NaN tie. My round-one counterexample now returns `route == 3`; D10 ties still return NaN.

Opus F4 — constant adjustment factor (was minor)

The pre-fit singleton guard is gone and replaced by a `constant_factor_terms` widening of the alias reconciliation. I tested both directions:

case	result
<code>fland_mode</code> constant at reference level	fits; , SE match R oracle < 1e-7
<code>civst_mode</code> constant at a middle level	fits; match < 1e-7
<code>fland_mode</code> and <code>civst_mode</code> constant simultaneously	fits; match < 1e-7
<code>fland_mode</code> a deterministic function of <code>civst_mode</code> (both varying)	still <code>not_estimable</code>
<code>hypertens := dm</code> (duplicated numeric covariate)	still <code>not_estimable</code>

So the widening did not weaken the rank-deficiency guard. **But** it left a residual, which is finding **F5** below.

Codex C1/C2/C3 — provenance

I re-attacked all three rather than reading the author’s tests as evidence.

- **C1** — `verify_evidence` now requires `set(entry) == {"id"}` and resolves it through the externally anchored registry. Rejected in my run: a v001-style `{path,`

sha256} entry, an unknown registry id, an {id, path} entry, and a registry-accepted file mutated after registration (**Corrupt result digest**). A **genuine** registry-accepted **no_result** is still honoured and never retried under either policy: **activation=false**, **terminal=not_run**, **coverage_state=None**, all 66 Atom-1 cells **not_run** with blank numerics — an operational terminal, not a scientific null.

- **C2** — **launcher.py** validates the whole-bundle digest against host-pinned external registry bytes before importing anything. Beyond the six files the author tests, I tampered **analysis.py**, **construction.py**, **interfaces.py**, **provenance.py**, **README.md** and **requirements-tested.txt**: every launch exited 2 with “*Execution bundle differs from controller-approved digest*” and created **no output directory**. A stray extra file changes the digest; a symlink is rejected explicitly; a planted **__pycache__/x.pyc** is rejected explicitly; restoring the bundle returns exit 0. Without the anchor, both the launcher and a bare **run.py** reject before outputs, while **run.py --help** still exits 0.
- **C3** — **select_slots** and **render** both route through **provenance.read_result**. The 999.0 estimate mutation, altered adjudication prose and a resealed envelope omitting the selected CSV are all rejected. My remaining concern is only the *provenance-consistency* gap recorded as **F6**.

2. Per-Atom coverage

Governed copied-layout CLI, run through the external launcher from a separate action bundle against my World copy, output outside the World, one invocation per Atom:

Atom	coverage rows	estimate rows	state	every cell
atom-000001 Population mortality	225	66	boundary	unresolved_source
atom-000002 Sibling comparison	8	6	boundary	unresolved_source
atom-000003 Surgical subgroups	52	45	boundary	unresolved_source
atom-000004 Route / regimen / initiation	36	36	boundary	unresolved_source
atom-000005 Cause-specific mortality	30	27	boundary	unresolved_source

No numeric field is populated in any of them. Under **--dispatch-policy direct** with definitions {D01: resolved, D08: adjustment set = A_P, D12: resolved, D14: resolved, resolved: true}, a hashed **source_evidence** item and **disclosure: true**, atoms 1 and 5 are still **boundary** with every cell **unresolved_source** — metadata cannot amend a frozen stop.

On the synthetic fixture path each Atom numerically exercises its supported components. The counts are **numerically unchanged from round one**, which is the check that matters here: the F3 and F4 widenings did not silently move any target across a registered stop.

Atom	rows	available	reference	unresolved_source	operational_failure
atom-000001	225	154	12	59	0
atom-000002	8	4	0	4	0

Atom	rows available		reference	unresolved_source	operational_failure
atom-000003	52	17	2	33	0
atom-000004	36	21	3	12	0
atom-000005	30	18	3	9	0

Estimator fidelity re-anchored independently: `engine.R`'s `cox` branch is byte-identical to `v001` (`coxph, ties="efron", robust=FALSE, na.action=na.fail, summary(conf.int=.95)`, no cluster/sandwich). My own strict counting-process Efron partial-likelihood optimiser (risk set `tstart < t <= tstop`, Efron denominator correction, SciPy Nelder-Mead) reproduced the `m01` crude HR 0.87375280597817 to $|\Delta\text{HR}| = 1.5\text{e-}7$.

Manuscript interface: `select_slots` binds all **292** fixed slots on the governed boundary tree (`{unresolved_source: 286, prose: 6}`), on the partial synthetic tree (`{available: 163, unresolved_source: 103, reference: 20, prose: 6}`) and on the author's unactivated tree (`{not_run: 168, unresolved_source: 58, available: 52, reference: 8, prose: 6}`, single shared `input_identity`).

3. The two findings

F5 (minor) — a constant *non-factor numeric* source covariate still stops a fit that `coxph` completes

`models.py:174-186` builds both `absent_terms` and the new `constant_terms` by iterating `factors.items()`, i.e. only `FACTOR_LEVELS` covariates. But `A_R` (`models.py:11`) and `A_CV` (`models.py:13`) also carry seven plain numeric 0/1 covariates — `dm`, `hypertens`, `afli`, `valv`, `hf`, `statin`, `many_diag` — which never enter `factors`. When one of those is constant in the risk set, its column is all-zero or all-one, `R` aliases it, and `models.py:186` fires.

Executed, on the package's own `code_calendar_pre2003` fixture risk set:

- `Engine.fit(..., x.assign(dm=0), ...)` → `Boundary("not_estimable", "Source Cox design is rank deficient...", "m112")`, `audit` `aliased=["dm"]`, `constant_factor_terms=[]`. Same for `statin=1`.
- On the identical CSV, `R 4.1.2 / survival 3.8.3` reports `aliased: dm`, **no warning**, and `ht beta -0.0245343280, se 0.5379655873` — identical to **10 decimals** to the same model with `dm` dropped. For `statin=1`: `aliased: statin`, `beta -0.1012498181, se 0.5341880675`, again identical to the reduced model.

This is the same situation the author's own comment at `models.py:183-185` already accepts for factor levels, applied to a different column type. It is **not a regression** — `v001` behaved the same way for numeric covariates and my round-one `F4` was scoped to `FACTOR_LEVELS` — so it is a new finding. Impact is bounded: under `purpose=governed` no fit runs at all, so no currently obtainable number changes. It would matter for the four executable adjusted models `m112/m122/m132/m162`, which fit on restricted subsets (the pre-2003 and post-2003 calendar arms, the BMI-observed

gravdb-linked subgroup) where a constant `valv/hf/statin` is plausible. Direction is safe: a stop, never a fabricated value. Hence **minor**.

Fix. Scope it to exactly what I demonstrated — plain numeric covariate columns that are constant in the risk set. Before the alias test add `numeric_cols=[c for c in cols if c not in factors and c!="family_id"]` and `constant_terms |= {c for c in numeric_cols if x[c].nunique()<2}`, keeping them in `constant_factor_terms`. `family_id` must be excluded explicitly: `models.py:145` substitutes it for `strata(family_id)`, so it is a stratum that contributes no design column and can never be legitimately aliased; sweeping it in would widen the guard past the evidence. Leave `models.py:152-154` and the rank-deficiency stop itself unchanged. Two lines.

Regression test. For `(dm, 0)` and `(statin, 1)`: assert the exposure and SE equal an independent `Rscript coxph` omitting only that constant term (`places=7`) and that the name appears in `constant_factor_terms`. Keep `x.assign(hypertens=x.dm)` and a `fland_mode` collinear with `civst_mode` as explicit negative controls that must still raise, and assert `"family_id"` never appears in `constant_factor_terms`.

F6 (minor) — `select_slots` binds a manuscript interface across results from different runs

`interfaces.py:9-19` correctly calls `verify_world` and then `read_result` per contributing Atom (digest, material identity, Atom identity, approved implementation, required output membership, all aggregate hashes). That fully answers Codex C3. What it does not check is that the contributing results belong to the **same run**: `read_result` takes no `input_identity` argument, and neither `synthetic` nor `accepted` is inspected or returned.

Executed: a tree with `atom-000001` from the synthetic suite (`synthetic: true, accepted: false, input_identity 62a8c14f3145...`) and `atom-000002..5` from my governed CLI run (`synthetic: false, input_identity 87ffa73797ec...`). Every envelope hash verifies, so `select_slots` returned **all 292 slots** and reported **111 as available**, with 38 synthetic Atom-1 estimates among the selected rows and nothing in the returned DataFrame distinguishing the two runs.

Stated so this is not overread: `execution/README.md:414-418` explicitly discloses the synthetic half — *“This checks integrity of synthetic/unaccepted previews as well as governed outputs; it is not permission to insert a preview into a manuscript”* — and `render_figures.py:26` writes a `SYNTHETIC-ONLY.txt` marker when it consumes a synthetic Atom-1 result. I am **not** reporting the preview capability as undeclared. What is undeclared anywhere in the World is that results from two different input manifests may be combined into one manuscript binding; the author’s own mixed fixture (`tests-out/unactivated-slots`, Atom 1 `not_run` + Atoms 2-5 `partial`) correctly shares a single `input_identity`, which shows the intended invariant. Hence **minor**.

Fix. Capture `input_identity` (and `synthetic`) from each `read_result` and raise if more than one distinct `input_identity` is present; expose the shared values so the caller cannot lose them. Do not add an acceptance gate — host adjudication

stays where the README puts it — and do not forbid synthetic previews, which the existing tests rely on.

Regression test. A tree mixing one synthetic Atom with governed Atoms must be rejected; the governed, partial and unactivated trees must still return 292 and now also report their single shared `input_identity` and `synthetic` flag. This fix breaks no passing test: every results tree the author's suite feeds to `select_slots` is already single-identity (`tests-out`, `slot-seal-tests` and `unactivated-slots` all 62a8c14f3145...; `boundary-slots` 87ffa73797ec...; my governed CLI tree 87ffa73797ec...).

4. What remains unavailable, and why

Plainly: **zero empirical numerical results are obtainable from this package**, and that is a property of the frozen source, not of the implementation.

- D01/D02 stop every Atom before any person row is read. D01 is genuine: R75's `as.IDate(1995-01-01)` is unquoted arithmetic. So nothing in Tables 1–3, Figure 1 or supplementary Figure 1 is currently reproducible.
- Target-level stops remain genuine and per-target rather than blanket: D08 (the published adjusted models `m02/m12/m22/m32/m42/m52/m62/m72/m82/m102/m202/m212`), D09, D10, D11, D12 (`t2_bso_ge55`, both `code_hyst` blocks), D13 (`code_smoking` crude/adjusted), D14 (`t3_other` adjusted), D15, D18/D08 (the 14 Schoenfeld panels), D19 (the 29 eTable 2 product denominators). Code-only siblings `m111/m112/m121/m122/m132/m161/m162` stay executable.
- The R319 cancer object is now built, but its input is itself blocked: with no historical cancer export, `clinical_dates` returns `Boundary("unavailable_input", "D05 historical cancer export absent for R319 split")`. The fix restores the source interval structure; it does not make the split computable on real data.
- Pinned R 4.4.1 / survival 3.7-0 / data.table 1.16.2 is unavailable; everything numerical here ran on R 4.1.2 / survival 3.8.3 / data.table 1.18.2.1 and is explicitly synthetic. `Engine(..., synthetic=False)` correctly refuses the installed engine, so the pin gate is verified and the pinned engine is not. `survminer/ggcoxzph` absent; Helvetica Neue absent.
- No confidential Danish/APTO row was mounted or read, no author-export equivalence was verified, no disclosure clearance exists, no old empirical session and no QGPU were used.
- Controller deployment is not established and the author correctly does not claim it. The trust anchor is two environment variables; anyone controlling the process environment can mint a self-consistent registry for a bundle they modified. That is the documented deployment boundary, and the design is still sound because a downstream host checking against the real registry rejects the unapproved `implementation_identity` that `run.py` stamps into every result. I record it as a limit, not a finding.

5. Why accept

All four of my round-one findings and all three of Codex’s are fixed, and I confirmed each with my own oracles rather than the author’s tests. The two remaining findings are minor, safe-direction, two-line fixes with named regression tests; neither withholds nor fabricates any currently obtainable number, and neither requires touching the frozen protocol, roster, deviation ledger, catalogue or manuscript. **accept** means faithful tested engineering within the explicit limitation that no empirical result exists — never observed science, original-protocol approval or replication.

Proof artifacts are under `scratch/r02/opus/` (`mytests/`, `tests-out/`, `cli/`, `trust/`, `trust2/`, `slots/`, `action/`, `copy-world/`, `nob-world/`). Nothing in `v002/world`, the immutable package, another reviewer’s outputs, controller files or repository resources was written: `sha256sum -c` still passes 78/78 and `find v002 -newermt <session start>` is empty.

Independence disclosure: my first scope check excluded only my own scratch path, so its output incidentally listed *filenames* under `scratch/r02/codex/`. I opened none of them, and no observation, finding, test or conclusion in this review derives from the other reviewer’s round-two work. The corrected check, which excludes all of `scratch/r02/`, shows the only files I created outside my own scratch are `reviews/r02/opus.yaml` and `reviews/r02/opus.md`.

r03 / codex.md

Codex engineering review — round 3

Verdict: **revise**. The round-three changes repair the defects they name, and the complete advertised suite passes, but three source-fidelity defects remain in supported lower-stage execution. These are narrow validation/landmark fixes, not requests for new experiments or protocol changes. This review does not accept or reject the frozen scientific design and reports no observed science.

What I inspected

I read the complete five Atom protocols and result regions, the full Molecule policy, shared construction and source-deviation records, target catalog/inventory, source provenance and schema/clinical mappings, all 969 lines of the supplied R program, the complete local paper and supplement, execution bundle, manuscript slots, and both figure bindings. The original package remains byte-identical outside the added `prereg/execution` and `prereg/engineering` directories: `diff -qr --exclude=execution --exclude=engineering` returned no differences, and all 78 entries in `prereg/CHECKSUMS.sha256` verified.

I did not read the other reviewer’s round-three output. I did read the earlier-round findings and the round-three author response. The new exact-byte acceptance binding, mandatory dataset SHA-256, binary checks, contiguous-calendar check, constant-numeric alias handling, and single-input slot rule are present in the code rather than merely asserted in the response.

Findings

C3-01 — major: longitudinal source states can reverse

The common input validator checks interval geometry at `prereg/execution/construction.py:341-356`. `prepare_target` checks binary `ht` and forbids `ht` from returning to zero at `prereg/execution/models.py:62-67`, but it does not enforce the corresponding source `tdc` and `cumtdc` invariants for other fields. It then selects current-zero contraindication rows at lines 68-70. That is insufficient for the registered “censor at first contraindication” operation in `atoms/atom-000001.md:164-168`.

My minimal woman has three contiguous intervals and `cvd=[0,1,0]`. The implementation accepted intervals beginning at 45 and 55, totaling 10 person-years. The registered pre-contraindication risk set is only the first interval, 5 person-years. A second case accepted `ht_sum=[6.0,0.5]` and moved duration category from 4 back to 1, despite the cumulative construction at `prereg/data-contract.md:51`. A third accepted `ht=[0,1]` with route `[1,0]`, classifying pre-use time as oral and exposed time as never although `exposure_intervals` requires the opposite at `prereg/execution/construction.py:368-400`. These histories are impossible outputs of the registered source construction, so accepting them is not a robustness choice.

Fix: reject, rather than repair, reversals in each source-persistent indicator used by a target; enforce finite nondecreasing `ht_sum` and `ht == (ht_sum>0)`; require route/regimen to agree with never/ever status; and apply analogous target-scoped consistency to cessation/surgery fields. Preserve independent siblings when an unused optional field is invalid. Add the falsifiers as regressions.

C3-02 — major: contiguous age intervals do not authenticate calendar time

The round-three fix correctly rejects internal age gaps. Its test at `prereg/execution/run_tests.py:257-267` does not check the supplied calendar values, however. `source_calendar` computes `year_start` from entry decimal time plus cumulative risk time at `prereg/execution/construction.py:306-321`; the model consumer merely requires `year_start` and filters it at 2003 at `prereg/execution/models.py:83-85`.

My contiguous `[45,50]` and `[50,55]` intervals supplied `year_start=[2000,2010]`. Both calendar branches accepted those values and put the second interval post-2003. The implementation’s own source kernel independently computes `[2000,2005]`, which keeps that interval in the pre-2003 R633 risk set. This is unverified required metadata capable of changing person-time, events and all four `m111/m112/m121/m122` outputs.

Fix: require a finite conserved clock within person (`next year_start = prior year_start + prior risktime`, at the registered precision) and bind its first value to entry decimal time when available. If the origin cannot be verified, return a calendar-input boundary. Do not invent a 2003 split, and keep D09 intact.

C3-03 — major: incomplete age-55 landmarks are marked available

The protocol explicitly says to assert one valid age-55 landmark row and stop Table 1 on failure (`atoms/atom-000001.md:136-143`; source R950). `Analysis.landmark` at `prereg/execution/analysis.py:96-101` only selects rows with `tstart==55`.

My fixture contains one correctly split `[45,55]/[55,60]` woman and one `[45,60]` woman. Both are at risk immediately after 55, but the second is silently omitted. The actual summary path at `prereg/execution/analysis.py:177-183` then marked landmark counts available as 1 exposed and 0 unexposed. This can affect both landmark-uptake windows and all 64 Table-1 slots.

Fix: if any interval spans age 55 without the exact split, stop all shared landmark/Table-1 consumers with an input boundary. Women genuinely censored at or before 55 should remain outside that landmark. Add the mixed split/unsplit case as a regression.

Per-Atom coverage

Atom	Executed synthetic coverage	Source-faithful limits and review disposition
atom-000001	225 coverage rows; supported synthetic model IDs <code>m01</code> , <code>m11</code> , <code>m21</code> , <code>m41</code> , <code>m111</code> , <code>m112</code> , <code>m121</code> , <code>m122</code> , <code>m132</code> , <code>m161</code> , <code>m162</code> ; full descriptive/flow/slot paths and both figures exercised.	Adjusted published fits, smoking HR, product denominator and PH remain explicitly blocked. C3-01, C3-02 and C3-03 require revision.
atom-000002	8 coverage rows; four supported summary rows available; full-person parent cleaning, eventual discordance, nonmultiplying join and all slots exercised.	Crude is correctly blocked by D11 and adjusted by D08. Shared cumulative-history validation in C3-01 also protects final-ever sibling selection.
atom-000003	52 coverage rows; crude <code>m71</code> and <code>m81</code> , surgery descriptions and all registered absent siblings exercised.	Older BSO and code hysterectomy-age targets remain D12; adjusted fits remain D08. Shared exposure/surgery persistence falls under C3-01.

Atom	Executed synthetic coverage	Source-faithful limits and review disposition
atom-000004	36 coverage rows; crude m51, m61, m101, initiation independence, route/regimen tie boundaries and all slots exercised.	D04/D10 remain scoped exposure-assignment limits and adjusted fits remain D08. Shared logical-history validation falls under C3-01.
atom-000005	30 coverage rows; crude m201, m211, m221, cause-specific censoring summaries/proportions and all slots exercised.	D09/D15 block an empirical horizon/classification; adjusted fits remain D08/D14. Duration histories inherit C3-01.

The synthetic suite is not an empirical oracle. Its available counts/models only show code-path reachability on authored fixtures.

Tests and actual outputs

The advertised command ran 38 tests in 102.926 seconds with zero failures/errors and reported `world_unchanged: true`, `synthetic: true`, and `pinned_engine_validated: false`. Its detailed log is [test-log.txt](#), and the report is [test-report.json](#).

My independently written falsifier is [adversarial_tests.py](#); its exact observed records are [adversarial-results.json](#). It completed successfully because all five asserted silent acceptances were reproduced, not because the reviewed behavior was correct.

I also copied the World and execution bundle to scratch and invoked the externally anchored launcher separately for all five Atom IDs with direct dispatch. Each produced a labelled synthetic `boundary` with no terminal. Selecting those actual outputs in explicit preview mode resolved all 292 unique manuscript slots with only `prose`, `unavailable_input`, and `unresolved_source` dispositions. Figure 1 and Figure 2 rendered from the actual Atom-1 output as one-page vector PDFs; text extraction retained all 17 and 14 numbered unavailable reasons respectively. These artifacts are under [copied-cli](#) and [copied-cli-figures](#).

Installed numerical software was Python 3.12.4, NumPy 1.26.4, pandas 2.2.3, SciPy 1.15.2, statsmodels 0.14.4, PyYAML 6.0.2, Matplotlib 3.10.0 and pypdf 6.11.0; R was 4.1.2 with survival 3.8.3 and jsonlite 2.0.0. The suite's Efron PHReg comparison, direct R Cox comparisons, tmerge boundary cases and model-based (not clustered) sibling-strata covariance all passed. They do not validate the unavailable frozen R 4.4.1/survival 3.7-0 combination.

Remaining source limits

No empirical numerical result is currently available. D01/D02 stop every governed target before person-row access because entry/time-zero, birthday boundaries, migration and censoring are unresolved. The narrower target stops D03-D19 remain explicit, including D08 adjusted attribution, D09/D15 cause/calendar scope, D10 type ties, D11 sibling crude strata, D12 surgery timing, D13/D14 wrong model objects, D18 PH inputs and D19 product denominator. Those are acceptable faithful boundaries; they must not be presented as null, equivalence, agreement or causal evidence.

No Danish/APTO data or old empirical session was used. Exact frozen R and survminer/ggcoxzph software were unavailable, and Helvetica Neue was unavailable for publication typography; the figure check used the documented preview fallback. RD, staggered DiD and matching are not registered estimators in this survival-study package and were not substituted.

r03 / opus.md

Round-three independent engineering review — Opus

Verdict: accept. Faithful, tested pre-data engineering within explicit, plainly-stated limitations. This is **not** scientific acceptance, not observed science, not replication, and not approval of the frozen protocol.

The single most important fact: *no numerical scientific result of any kind is available from this bundle.* In governed mode the shared D01/D02 gate stops every requested component of all five Atoms before any person row is read. My five governed runs produced 351 coverage rows, every one **unresolved_source**, and every numerical field in **estimates.csv**, **summaries.csv**, **flow.csv**, **characteristics.csv** and **products.csv** blank. **source-comparison.csv** is correctly an empty declared interface in every Atom, and no published value is copied anywhere into a result.

Reviewed World v003/world; reviewer scratch **scratch/r03/opus**. The reviewed World, the original package, controller state and the other reviewer's outputs were never written. I verified this at the end: my working copy is byte-identical to v003/world across all 95 files.

How I reviewed

I deliberately did not treat the author's documentation or the author's test suite as oracles. I read the frozen R source directly (**mikkelsen-supplement2.txt**, sha256 **a386ae04...0cfeea**, matching **source-provenance.yaml**) at R56–65, 110–142, 143–175, 196–262, 340–448, 482–530, 566–630, 630–672, 694–730, 738–800, 804–832 and 940–968, and then built five **reviewer-owned literal-R oracles** in **data.table** 1.18.2.1 plus one **independently coded Efron Cox** written from the partial-likelihood definition with analytic score and information, importing neither **survival** nor **lifelines**.

I also read all four prior review YAMs and the author’s response table, and re-tested every prior finding rather than assuming either earlier reviewer was right.

Ran the advertised suite from a copied World at a path the author never used: **38 tests, 0 failures, 0 errors, world_unchanged: true, pinned_engine_validated: false.**

Per-Atom coverage

Catalog cross-check: 102 targets (21 estimates, 32 characteristic, 29 product, 12 summary, 6 context, 1 flow, 1 diagnostic). **Zero** catalog estimate targets are missing from `models.SPECS` and zero SPECS entries are absent from the catalog. Exported rows equal exposures $\times$ components exactly for every Atom.

atom-000001 — Population mortality (66 estimate rows, 17 flow, 64 characteristic, 29 product, 14 panel)

Implemented and independently verified: exclusion union counted separately from its seven overlapping reasons; last-`tstop` versus `tstart==55` landmark correctly distinguished (R806 vs R950); binary m01, duration m11, contraindication m21, cessation m41, code-calendar m111/m112 + m121/m122 (adjusted correctly omitting `year_cat`, R640/R654), surgery-adjusted m132 sharing the main crude (R667), BMI m161/m162, smoking summary only.

My checks: a 24-woman hand-computed fixture reproduced `t2_primary ht=0` events 12 / person-years 294.0 / IR 408.163265 exactly, correctly withheld the 2-event exposed group as **suppressed**, and matched flow never/ever 12/12, duration nodes 0/4/0/4/4, follow-up q25/median/q75 14.5/16.0/17.75 and end-age median 61.0 (R type-7 quantiles). The literal R136 oracle over five episode layouts — including a nested pair whose later episode ends earlier, a three-episode case whose largest `htstop` is in the middle, and a tie on max `htstop` — matched `construction.cessation` to $<1e-9$ (55.246 / 51.246 / 52.246 / 50.646 / 56.246), confirming the round-one F2 repair is real. R486 contraindication censoring retained exactly the 5.0 pre-event person-years. Duration cuts matched R at every break.

Correctly blocked: adjusted m02/m12/m22/m42 (D08), smoking HR (D13), all 29 S2 product proportions (D19), all 14 Schoenfeld panels (D08/D18) — and the panels stay blocked in synthetic runs too, which is right.

atom-000002 — Sibling mortality (6 estimate rows)

A literal R45–47 + R505 oracle retained exactly {SYN-01, SYN-02, SYN-07, SYN-08, SYN-09} from six constructed families; `construction.siblings` matched row for row. Distinct fathers accepted, missing mothers dropped, singletons and concordant families dropped, **all** sisters of a discordant family retained with their full contemporaneous histories, and the many-to-one join did not multiply rows (10 in, 10 out). Selection on *eventual* discordance is preserved as source-defined selection and is not recast as baseline randomisation. Crude m31 correctly stopped at D11 with

no strata added; adjusted m32 retains `strata(family_id)` in the extracted formula and is stopped at D08. The engine is model-based, never a family-clustered sandwich (`robust=FALSE`, engine.R:45).

atom-000003 — Surgical mortality (45 estimate rows, 2 summary targets)

`t2_hyst` and postoperative `t2_bso_45_54` implemented with their own cancer censoring and overlapping memberships. I checked the R583 boundary precisely: `bso_age` 54.999 is retained, 55.0 and 55.001 are excluded, matching R's `bso_age>=55 → bso:=2` overwrite. The BSO death-age target uses each decedent's `tstop` conditional on dying (R597–599) and is not confused with population N. `t2_bso_ge55`, `code_hyst_lt55` and `code_hyst_ge55` retain the exact D12 stops — no speculative postoperative filter is run — while the younger-group outputs survive independently, exactly as the Atom rationale requires.

atom-000004 — Treatment heterogeneity (36 estimate rows)

A literal R382–427 `data.table` oracle matched `exposure_intervals` on every row of four persons: a route flip, an oral-vs-transdermal tie (both NaN → D10 stop, **no** `nafill` carry), an unmapped NASSPRO form routed to “other” per R394 (the round-one F3 repair), and a Mirena-driven monophasic regimen. A literal R618 oracle matched `initiation` codes for first exposure at 45.05, 51.999, 52.0 and 57.0. The overwritten R616 recipe is correctly not executed as a second sensitivity, and no final route is projected backwards. Initiation remains computable when route/regimen are tied — the promised independent component disposition holds.

atom-000005 — Cause-specific mortality (27 estimate rows)

Cause code 12 is classified “other” and never silently added to cancer. An A-02 + list-49 23 conflict and an all-missing record both raise D15; an unclassified death raises “Unclassified death is not other mortality” rather than defaulting to other. Competing deaths censor each cause-specific fit with event 0 at their own `tstop` with risk time preserved — no risk subtraction, no Fine–Gray, no cumulative incidence. Cause proportions use the common classified-death denominator (R797–799). Adjusted cardiovascular is blocked pending the R756-vs-footnote reconciliation, and adjusted other-cause remains the absolute D14 stop: neither m92 copying nor m222 substitution is reachable from the adapter.

Estimator and inferential fidelity

My independently coded Efron partial likelihood (analytic score/information, no survival library) matched the native engine on 1319 intervals / 135 events:

model	abs(Δ)	abs(Δ_{se})
~ ht	2.8e-16	5.6e-17
~ ht + dm	5.5e-14	2.2e-15
~ ht_cat2 (3-level factor)	1.4e-11	5.5e-13

HR, lower and upper matched to eight decimals. This confirms Efron ties, **model-based** (not sandwich, not clustered) variance and the native 95% CI. A grep of the delivered *.py/*.R finds no weights, matching, propensity, frailty, robust, RD, DiD or competing-risk estimator anywhere in the analysis path; `statsmodels.PHReg` appears only inside `run_tests.py` as an oracle. Reference rows carry HR 1 explicitly labelled **reference** with **no** fabricated CI, and `present_estimates` rounds only the presentation layer while the machine table keeps full precision. D16’s missing→1 repair exists nowhere.

The F5 fix behaves: with constant nuisance covariates (`dm=0`, `statin=1`, `valv=1`, `hf=0`, `hyst=0`, `bso=1`) the full source formula is retained, the aliased constants are recorded in `constant_numeric_terms`, and the exposure HR is exported (0.86228979, 0.55473047–1.34036928); a genuinely collinear varying pair still stops with “Source Cox design is rank deficient”. `family_id` is excluded from the allowed alias set by construction, so a stratum can never be laundered as a harmless alias.

Dispatch, evidence and registration identity

Direct dispatch bypasses **only** outer activation. It could not clear a single D-code stop, and manifest `definitions: {"D01": "resolved", "resolved": true, "amendment": "approved"}` with `source_evidence` claiming `"clears": ["D01", "D02", "D08"]` changed nothing. Forged predecessors were rejected two ways (“Predecessor lacks controller acceptance”; “Predecessor must reference a controller accepted-result ID”) before any output existed. A host-accepted self `no_result` was never retried under **either** policy, and produced the full 14-file interface with `not_run` and blank numbers — not a fabricated scientific state. `native.py`’s parser/evaluator are byte-identical to the repository originals, so strong Kleene semantics (NOT unknown is unknown) are the registered ones.

Registration identity, 10/10 as required:

edit	expected	observed
Atom narrative body	accept	accept (material unchanged, realization changed)
<code>atom_result: complete</code>	accept	accept
<code>+ registered_state: partial</code>		

edit	expected	observed
Molecule body	accept	accept
invalid	reject	“Invalid realized Atom state”
registered_state key		
pending Atom carrying a state	reject	“Pending Atom cannot carry a realized state”
protocol cohort text	reject	“Atom registration altered”
activation expression	reject	“Atom registration altered”
Molecule membership	reject	“Molecule registration or membership altered”
Molecule registration field	reject	“Molecule registration or membership altered”
frozen	reject	“Frozen material altered”
data-contract.md		

Bundle identity of my **relocated** copy equals the reported digest 92459bc6ebbe979f09f0775db64aa5841a1f8f926f8e54946e0eb1b9591b64e3, i.e. it is path-independent. A bundle with the in-bundle verifier neutralised **and** the D08 stop deleted was rejected by the protected external launcher before import; a corrupted registration lock, a symlink, a .pyc, a substituted registry under the original host pin, an in-bundle registry and a missing anchor all failed closed.

Prior findings: re-tested, not assumed

finding	status under my own falsifying cases
r01 F1 — R319 cancer missing from the tdc allowlist	fixed (engine.R:18-21)
r01 F2 — cessation fallback by last htstart	fixed; literal R136 oracle, 5/5
r01 F3 — unmapped dose form nullified route	fixed; literal R394 oracle
r01 F4 / r02 F5 — legal constant covariate stopped a fit	fixed; native + independent oracle
r02 F6 — slots mixed results from different runs	fixed; both mixed trees rejected
r01 codex 1 — caller-authored predecessor evidence	fixed; both forgery shapes rejected
r01 codex 2 — no immutable anchor for the code	fixed; external launcher + digest

finding	status under my own falsifying cases
r01/r02 codex — resealed aggregates reaching manuscript/figure	fixed (C2-01); reseal rejected canonically, accepted only in explicit preview
r02 codex C2-02 — dataset descriptors without content hashes	fixed; missing/short/non-string all rejected before any output dir
r02 codex C2-03 — indicator domains and calendar continuity	fixed and correctly scoped (see O2)

The v002→v003 diff is purely additive: no supported model, kernel or output was removed.

Figures and manuscript interface

All 292 fixed slots resolve. Both figures export at exactly **130.70 mm**. Visually inspected at manuscript width: boundary figure 1 prints “Unavailable [1]...[17]” with numbered explicit reasons and never a fabricated count; figure 2 renders all 14 S1 panels in the exact source term order (ht, year_cat, fland_mode, civst_mode, udda_mode, q_ind_mode, parity_cat, dm, hypertens, affi, valv, hf, statin, many_diag) as blocked with D08/D18 reasons and no substitute smoother. A `flow.csv` row carrying `n=999` on an `unresolved_source` status was rejected: “eligible: unavailable counts must not contain values”. Canonical consumption requires externally bound result bytes; preview writes `PREVIEW-ONLY.txt` and `figure-provenance.json` and licenses nothing.

Finding (one, low severity, not blocking acceptance)

O1 — `estimates.csv` `n` overlaps across contemporaneous exposure groups. `analysis.py:119` counts distinct women per exposure group; because ever-use is persistent (R427), every eventually-exposed woman also contributes to group 0, so in my 24-woman fixture the two groups reported `n=24` and `n=12`. The source tables have no `n` column at all (R491/R513/R533/R570/R585 aggregate only `event` and `risktime`), and no slot in `result-slots.yaml` binds `n` from `estimates.csv` (`run_tests.py:623` pins the same shape), so nothing reaches the manuscript. Fix: one clarifying clause in the README `estimates.csv` paragraph; regression: assert the overlap explicitly in the existing hand-oracle test.

Examined and accepted (not defects)

Contiguity scoping. `contiguous=True` is requested only from `source_calendar` (`construction.py:308`) and the `code_calendar_*` preparation (`models.py:66`); I confirmed both reject a gapped history and that a contiguous one yields `year_start`

2000.0/2005.0. `analysis.py:181` (`landmark_uptake`, R957–960) and the `year_cat` Table 1 cells (`analysis.py:26`) consume the same clock from a caller-supplied column without that check. Not a defect: the README’s scope statement is literally accurate, the input contract already assigns calendar conservation of a supplied column to the caller, and D01/D02 makes the path unreachable in the frozen World. Passing `contiguous=True` on those two consumer paths would be a zero-science tightening if the author wants the symmetry later.

Also examined: missing versus failing R engine (boundary vs operational terminal — deliberate, tested, consistent with the Atom verdict vocabulary); disclosure suppression stricter than R729’s `event:=NA` (withholds recoverable PY/rate, documented, status-labelled, nothing invented); governed mode never loading declared datasets (documented substantive limitation, not an unfinished wrapper); bare `python run.py` on a tampered bundle (documented — the external launcher is the trust boundary, and it rejects); the frozen catalog’s “Year of 55th birthday:” group label on `t1_31–t1_39` (immutable original package, slot keys still unique, not mine to change).

What acceptance does and does not mean

It means: the reference code implements every source-supported component of all five Atoms, its kernels reproduce the frozen R exactly where I could build an oracle, its estimator is the source’s `coxph` with Efron ties and model-based variance, its stops are real registered source gates that no caller input can clear, and its identity, evidence and figure interfaces resist the forgeries I could construct.

It does not mean any number was produced. Every cell of Table 1, Table 2, Table 3, Figure 1, supplementary Figure 1 and supplementary Table 2 remains unavailable, because D01/D02 leave cohort time-zero, age boundaries and migration censoring genuinely unresolved between P2–3 and R39–77/R279–297, and D03–D20 gate the rest. Producing any of them requires a disclosed pre-execution amendment under `governance.md`, not a code workaround.

Complete registered Molecule

```
id: molecule-000001
molecule_mode: prereg
molecule_shape: convergent
topic: MHT mortality replication
rating: C
draft: false
membership:
- atom: atom-000001
  role: primary
  description: This is the paper's central mortality question. Cohort flow,
    landmark characteristics and utilization descriptions orient it; duration,
    contraindication/cessation sensitivities, PH assessment and
    ↪ source-code-only
    calendar/BMI/surgery-adjusted checks interrogate that same association.
    They are not separate experiments.
- atom: atom-000002
```

role: secondary
 description: The sibling comparison changes the identifying comparison to families selected for discordant exposure. Its judgment about shared
 ↪ familial
 confounding is interpretable independently from the nationwide
 ↪ association;
 it warrants its own Atom, while its crude/adjusted variants and
 ↪ descriptive
 size remain together.

- atom: atom-000003
 role: secondary
 description: The surgery analyses address a clinically distinct population and timing of ovarian removal. Hysterectomy, BSO45-54 and BSO>=55 are related contrasts within that question, with overlapping memberships and separately preserved risk sets/results. Age-stratified hysterectomy code variants remain in this module; an unsupported older group cannot erase supported younger-group evidence.

- atom: atom-000004
 role: secondary
 description: These related exposure contrasts address heterogeneity in the MHT association. They stay together because they partition the same
 ↪ treatment
 history for one source question, with distinct coefficients and boundaries rather than a composite effect. This is independent of surgery-specific population selection.

- atom: atom-000005
 role: secondary
 description: Cardiovascular, cancer and other mortality jointly elaborate the all-cause association while sharing the cause-register horizon and classification. Their distinct cause-specific estimates and missing
 ↪ components
 are preserved within one original analysis module; they are not one Atom per cause or coefficient.

literature: []

frontier:
 rule: Faithfully extract the published MHT mortality question and all
 ↪ original
 targets.
 avoid: No new questions, speculative repairs, extra analyses or
 ↪ causal/replication-success
 claims.
 directions: []

preregistration:
 mode: canonical
 question: What are the source-defined associations between systemic
 ↪ menopausal
 hormone therapy and long-term mortality in the original Danish cohort and its specified comparisons?
 scope: Original1950-1977 female cohort, age45 entry, all-cause
 ↪ through2023-07-31
 and cause-specific through2022-12-31, original source analyses only;
 ↪ unresolved
 inputs/rules remain explicit no-estimate boundaries.

storyline: Reconstruct the published cohort flow, follow-up, uptake and age55 characteristics before describing the nationwide binary and
 ↪ duration-specific
 all-cause mortality associations. Report the contraindication sensitivity, then the independently interpretable sibling contrast, then cessation sensitivity in the source Table2 order. Describe hysterectomy and BSO

age groups with their separately preserved risk sets and death-age
 ↳ summaries;
 next present treatment form, regimen and initiation age. Finally report
 cardiovascular, cancer and other mortality at the shorter cause-register
 horizon. The primary diagnostic and supplementary product-use table remain
 attached to population mortality; code-only source sensitivities appear
 in a clearly marked appendix. Every source result is inventoried even
 if unavailable. No favorable conclusion is prewritten. Each estimate and
 interval, model provenance and boundary governs its own wording. All
 ↳ modules
 are canonically active TRUE; their shared data-construction dependencies
 are checked inside protocols, not manufactured scientific activation
 ↳ edges.
 Sibling/surgery/exposure/cause evidence informs interpretation of
 ↳ population
 mortality but does not supply its data or repair an unresolved primary
 model.
 interpretation:
 - kind: restriction
 condition: 'TRUE'
 says: For every component, unresolved specification, unavailable input,
 suppression, nonestimation or operational failure prevents a numerical
 claim for that component. Never copy publication values into replication
 results, choose a conflicting model, infer null from a boundary, or
 equate coverage with agreement. Apply this restriction before all
 ↳ permissions.
 - kind: restriction
 condition: state(atom-000001) == "boundary"
 says: 'Population mortality: no source-faithful empirical component is
 available; state exact affected targets and causes. Other Atoms retain
 their own scope.'
 - kind: primary
 condition: (state(atom-000001) == "complete") OR (state(atom-000001) ==
 "partial")
 says: 'Population mortality: describe each available source-target
 ↳ estimate
 with its 95% interval, descriptive quantities and availability; retain
 discordant signs/precision. For partial coverage explicitly list omitted
 numerical components and why. Do not summarize this coverage state as
 a scientific verdict.'
 - kind: restriction
 condition: state(atom-000002) == "boundary"
 says: 'Sibling mortality: no source-faithful empirical component is
 ↳ available;
 state exact affected targets and causes. Other Atoms retain their own
 scope.'
 - kind: supplementary
 condition: (state(atom-000002) == "complete") OR (state(atom-000002) ==
 "partial")
 says: 'Sibling mortality: describe each available source-target estimate
 with its 95% interval, descriptive quantities and availability; retain
 discordant signs/precision. For partial coverage explicitly list omitted
 numerical components and why. Do not summarize this coverage state as
 a scientific verdict.'
 - kind: restriction
 condition: state(atom-000003) == "boundary"
 says: 'Surgical subgroup mortality: no source-faithful empirical component
 is available; state exact affected targets and causes. Other Atoms
 ↳ retain
 their own scope.'

- kind: supplementary
condition: (state(atom-000003) == "complete") OR (state(atom-000003) == "partial")
says: 'Surgical subgroup mortality: describe each available source-target estimate with its 95% interval, descriptive quantities and availability; retain discordant signs/precision. For partial coverage explicitly list omitted numerical components and why. Do not summarize this coverage state as a scientific verdict.'
- kind: restriction
condition: state(atom-000004) == "boundary"
says: 'Treatment heterogeneity: no source-faithful empirical component is available; state exact affected targets and causes. Other Atoms
→ retain
their own scope.'
- kind: supplementary
condition: (state(atom-000004) == "complete") OR (state(atom-000004) == "partial")
says: 'Treatment heterogeneity: describe each available source-target estimate with its 95% interval, descriptive quantities and availability; retain discordant signs/precision. For partial coverage explicitly list omitted numerical components and why. Do not summarize this coverage state as a scientific verdict.'
- kind: restriction
condition: state(atom-000005) == "boundary"
says: 'Cause specific mortality: no source-faithful empirical component is available; state exact affected targets and causes. Other Atoms
→ retain
their own scope.'
- kind: supplementary
condition: (state(atom-000005) == "complete") OR (state(atom-000005) == "partial")
says: 'Cause specific mortality: describe each available source-target estimate with its 95% interval, descriptive quantities and availability; retain discordant signs/precision. For partial coverage explicitly list omitted numerical components and why. Do not summarize this coverage state as a scientific verdict.'
- kind: supplementary
condition: done(atom-000001) AND done(atom-000002)
says: Read population and sibling evidence together as different selected populations and comparison structures; report estimates/intervals
→ separately,
with no formal difference test, pooled effect or equivalence rule. A family comparison neither proves removal of all confounding nor repairs a blocked population adjusted result.
- kind: supplementary
condition: done(atom-000003) AND done(atom-000004) AND done(atom-000005)
says: Read surgery, treatment and cause evidence as source-specific scope qualifications. Overlapping subgroups and distinct cause hazards cannot be averaged into a single score; one available result cannot stand for an unresolved older-BSO or other-cause adjusted target.
- kind: residual
condition: 'TRUE'
says: All remaining pending, unrun, operationally failed, inconclusive or incompletely classified scope remains explicit. If no evidence
→ exists,
state that no replication finding is available. No original source or package completion alone establishes replication agreement.

Authority is the complete membership and Molecule registration above and each Atom registration; shared referenced source documents are incorporated by their protocols.

Current interpretation

All five registered analysis modules are pending. No empirical results or realized evidence citations exist. The C compatibility rating does not claim acceptance, replication failure or scientific review. Stage1 is extraction with a source-fidelity self-check; Stage2 is transcription.

atom-000001: Population mortality

AUTHORITY AND SCOPE. Incorporate every applicable instruction in prereg/data-contract.md (schema table, ordered construction, inference and output contract), replication/clinical-codes.yaml and replication/deviations.md. These are part of this protocol, not optional implementation advice. P/S/R are the checked source locators defined in source-map.md. Only explicit source rules or loss-less engineering may be implemented. Each STOP_UNRESOLVED_SOURCE / STOP_UNAVAILABLE_INPUT emits a component-specific coverage row and absent numerical output. A shared upstream gap propagates to all and only its consumers. Do not treat this stage's documented gap as a realized Atom result: all Atoms remain prereg and pending.

COMMON POPULATION / TIME. Women born1950–1977, alive and resident at45; immigration before25; no prior source contraindication, MHT or bilateral oophorectomy. Unit is woman's (start,stop] age interval; endpoint is CPR death, censoring as P3, end2023-07-31 except cause-specific2022-12-31. Exact calendar/birthday and migration conflicts D01/D02, source-file mapping D03/D05 and applicable covariates D06/D07 must be resolved before consuming estimates. No missing key, date or excluded-history gap may be converted to zero/no-history. The source ever-use process starts unexposed and never returns to unexposed; source cumulative-duration process and its unresolved episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95% intervals and source software defaults in data-contract.md. P's two-sided .05 convention may describe individual source-specified tests; it does not adjudicate replication. No power calculation, multiplicity correction, alternative estimator, extra subgroup or empirical threshold is added. Use results/atom-000001/estimates.csv, summaries.csv, coverage.csv, adjudication.md and source-comparison.csv with the shared exact schemas. Do not create a current result file from published values. At runtime compute each source-defined output independently once its required inputs/rules are resolved; an adjusted fit's failure never overwrites crude/summary siblings. Retain exact target/exposure/component labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has terminal no_result and no registered local state. For an evidence-backed completed module, apply verdicts in the listed first-matching order. Boundary/partial/complete concern

coverage only. The scientific evidence is the entire labelled estimate/CI/descriptive record. Mixed signs and uncertain estimates remain mixed; missing results are not neutral. No local key means causal validity, significance, equivalence, replication success or independent review. Shared cohort joins and checks are inside this protocol, not separate experimental nodes.

SOURCE TARGETS. P Results pp4–6, Figure1, Table1, Table2 primary/duration/contraindication/as-treated blocks, discussion p8 follow-up and contraindication descriptions; S2 product distribution and S figure1; R279–299,429–480,485–500,525–542,632–667,698–725,806–967. Code-only checks are explicitly distinguished from published results. Original output order is target-inventory.md, not R execution order.

EXECUTABLE PSEUDOCODE. 1. **REQUIRE_SOURCE** for each output. Reconstruct only source-resolved logical inputs and shared cohort/intervals. Retain exclusion flags separately and union for Figure1. Emit flow.csv keyed by **node** with **n,status,reason** for eligible, excluded_total, thrombophilia, liver, arterial, venous, breast_gyn, previous_mht, previous_bso, included, never, ever, duration_lt1, duration_1_2_9, duration_3_4_9, duration_5_9_9, duration_ge10. Count eligibility before the exclusions; reasons overlap. Exposed and unexposed are end-of-follow-up groups for the flow, not baseline exposure groups in Cox. 2. Build last interval per person ordered descending tstop (R806); sum risktime per person and compute median,25th,75th percentiles (R817) in years, preserve source R quantile behavior once pinned. Emit total deaths/proportion, total end-of-follow-up ever-use/proportion, ever-use proportions for1950/1960/1970 birth cohorts, exposed cumulative-duration median/IQR and final duration counts/proportions, and end-age median in summaries.csv. End-age median P8 is a descriptive target using each woman’s final tstop, not a survival estimate. Death age is not life expectancy. If D01–D04 block their inputs, emit precise absent rows; no published counts enter results. 3. Landmark Table1: retain source intervals starting exactly55 (R950), group by past/present MHT at55. Assert one valid landmark row per included woman; otherwise stop Table1 rather than count duplicates. Report the exact characteristic roster in target-catalog.yaml: country including unknown, parity mean/SD and0/1–2/>=3, marriage status incl unknown, four education levels+unknown, imputed income q1–q4, calendar periods2002–08/2009–15/2016–23, six individually listed cardiometabolic indicators, >=3 contacts, prior hyst/BSO. Use characteristics.csv: **characteristic,category,exposure,n,percent,mean,sd,status,reason;** no t-tests (R953 test=F). Denominators are group-specific landmark N, unknowns explicit except source-disclosure income imputation. Missingness proportions from each woman’s final row R963–967 are separate scalar targets, not Table1 missingness estimates. 4. At55 calculate proportions for [2004,2007) and[2021,2024) as source R957–960, not all2007–2020 as a new trend. S2 requested cells are products.csv: **atc,regimen,form,percent,status,reason,** fixed source29-row roster. Source R430 groups nonmissing htac rows’ risktime and divides by total; P/S describe proportion of MHT. D19 stops the proportion target until the denominator is established; do not replace with dose or prescription counts. 5. Population main binary: summary by ht=0/1, events=sum(event), person_years=sum(risktime),

$IR=10000 \times \text{events} / \text{person_years}$. Crude m01: $\text{Surv}(\text{tstart}, \text{tstop}, \text{event}) \sim \text{ht}$. Adjusted m02 R461: $\sim \text{ht} + \text{year_cat} + \text{fland_mode} + \text{udda_mode} + \text{q_ind_mode} + \text{parity_cat} + \text{dm} + \text{hypertens} + \text{afli} + \text{valv} + \text{hf} + \text{statin} + \text{many_diag} + \text{civst_mode}$. Printed adjustment list is $A_P = \text{year_cat} + \text{fland_mode} + \text{udda_mode} + \text{q_ind_mode} + \text{parity_cat} + \text{dm} + \text{hypertens} + \text{afli} + \text{valv} + \text{hf} + \text{statin} + \text{many_diag}$. D08 prevents selecting or estimating an alternative for the published adjusted target. Retain crude/count siblings if their cohort/exposure inputs are fully resolved. No extra code-versus-paper fit is created. 6. Duration: 0, <1, 1–2.9, 3–4.9, 5–9.9, >=10 years, with the R cut bounds specified in shared construction. m11 $\sim \text{ht_cat}$; m12 $\sim \text{ht_cat} + A_R$. Reference0, each nonreference coefficient/95%CI, no trend test. D04/epsilon/upper-bound conflicts stop affected duration assignments; D08 separately stops adjusted attribution. 7. Contraindication sensitivity R486: retain pre-contraindication intervals where $\text{thromb} = \text{vte} = \text{cvd} = \text{breastgyn} = \text{liver} = 0$, thus censor at first qualifying contraindication. Summary/m21 $\sim \text{ht}$; m22 $\sim \text{ht} + A_R$. Source time-dependent boundary-day convention must match resolved tmerge semantics. Emit source R489 percentage of all deaths with preceding contraindication as separate descriptive target. Main risk set is not censored for these conditions. 8. As-treated R136/R526: source cessation recipe extends episode end by 90/365.2; compare to next episode start; mark gaps when $\text{htstop90} < \text{next start}$, also terminal episode, select earliest marked stop per person, round3 decimals. Before first MHT htstop90 is forced 0; retain intervals $\text{htstop90} == 0$. Summary/m41 $\sim \text{ht}$; m42 $\sim \text{ht} + A_R$. No reentry, no additional grace period, no drug-specific deviation. D04 episode and open-end ordering must be resolved; never use cessation sensitivity to replace the blocked ever-use target. 9. PH: S figure1 is 14 term panels in order $\text{ht}, \text{year_cat}, \text{fland_mode}, \text{civst_mode}, \text{udda_mode}, \text{q_ind_mode}, \text{parity_cat}, \text{dm}, \text{hypertens}, \text{afli}, \text{valv}, \text{hf}, \text{statin}, \text{many_}$ (visually checked). R465–466 uses $\text{cox.zph}(\text{m02})$, ggcoxzph . R480 computes $\text{cox.zph}(\text{m12})$ without a published plot; retain diagnostic existence in audit, do not add another manuscript figure. After D08/D18 resolved export approved $\text{schoenfeld_points.csv}(\text{term}, \text{time}, \text{residual})$, $\text{schoenfeld_curves.csv}(\text{term}, \text{time}, \text{fitted}, \text{lower}, \text{upper})$, $\text{schoenfeld_panels.csv}(\text{term}, \text{status}, \text{reason}, \text{p})$ and $\text{schoenfeld_ticks.csv}(\text{term}, \text{position}, \text{label})$ preserving the original transformed coordinates and age labels. All scientific transforms/smoothing/intervals come from the pinned original R plotting method; Python draws them only. Exact $\text{ggcoxzph}/\text{survminer}$ implementation/version is an input prerequisite, not freedom to choose a smoother. Report qualitative assessment and source P-values without inventing PH accept/reject rules or fitting a repaired model. 10. Code-only sensitivity rows (not published Table2 values): R633/647 interval-start split <2003 versus ≥ 2003 ; m111/m121 $\sim \text{ht}$, m112/m122 $\sim \text{ht} + (A_R \text{ without } \text{year_cat})$. D09 stops unresolved calendar risk sets; no2003 splitting repair. R663 m132 $\sim \text{ht} + A_R + \text{hyst} + \text{bso}$, shares original population crude/summary; source code explicitly supports these adjustments only for this separate specification. R699 BMI observed-interval subset, categories BMI <30 versus ≥ 30 , m161 $\sim \text{ht}$, m162 $\sim \text{ht} + A_R + \text{bmi_cat}$. BMI measurement source/timing D06 must resolve; never restrict to newly chosen measurements. Smoking R714/715 has separately supported smoking-complete summary, but m171/m172/rt172 conflict D13: no smoking HR or merged table estimate, no

repaired fit. Emit exact `code_*` rows in a source-code appendix interface, not as published numerical comparisons. All applicable shared source conflicts remain gates. 11. Assert component identities and comparator labels, round only presentation, apply source disclosure treatment (event counts<3 suppressed, not imputed). Export rows for every inventory cell. Report any combination of crude/adjusted/count/PH availability; the primary association remains unestimated if its specified model is unresolved even when sensitivities produce results.

SCIENTIFIC MEANING. Population associations depend on measured-confounding control and informative-treatment/censoring limitations. No causal benefit, clinical safety, equivalence or entire-population agreement follows from precision, an HR below1 or completed outputs. Source-only sensitivity results are supporting evidence within this question, never a replacement primary endpoint.

Complete registration fields

```
id: atom-000001
topic: Population mortality
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds: []
preregistration:
  version: 1
  mode: canonical
  question: What is the source-defined association of systemic MHT with
    ↪ all-cause
    mortality in the eligible nationwide cohort?
  rationale: This is the paper's central mortality question. Cohort flow,
    landmark characteristics and utilization descriptions orient it; duration,
    contraindication/cessation sensitivities, PH assessment and
    ↪ source-code-only
    calendar/BMI/surgery-adjusted checks interrogate that same association.
    They are not separate experiments.
  assumption: The source question is observational. Correct record linkage,
    source-equivalent histories, treatment/censoring specification and
    ↪ reported-model
    attribution are prerequisites; neither schema nor source agreement proves
    absence of confounding.
  target_strength: associational
  target_shape: association
  ratings:
    importance: 3
    novelty: 1
    feasibility: 1
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - VNDS
    - DOD
    - LMDB
    - LPR_ADM
    - LPR_DIAG
```

- LPR_OPR
- LPR_SKSOPR
- LPR_A_KONTAKT
- LPR_A_DIAGNOSE
- LPR_A_PROCREGISTRERING
- MFR
- CIV
- IND
- UDDA
- UDDF

activation: 'TRUE'

verdicts:

- key: boundary
 - region: After a completed evidence-backed assessment, no requested
 - ↪ empirical
 - component can be produced because every component is
 - ↪ source-underdetermined
 - or unavailable. Operational failure is excluded from this region.
 - says: Report precise no-estimate boundaries. There is no empirical
 - ↪ mortality
 - finding or replication-agreement conclusion.
- key: partial
 - region: At least one requested empirical component is available and at least one is explicitly unresolved, unavailable, not estimable or
 - ↪ disclosure-suppressed
 - after completed assessment. No operationally unfinished component is hidden by this state.
 - says: Report the available target-specific estimates, intervals and
 - ↪ descriptions
 - and all unavailable components. Partial coverage is not statistical confirmation or a uniform evidence direction.
- key: complete
 - region: All requested empirical components within this Atom, including code-only targets explicitly assigned to it, have resolved source
 - ↪ specifications
 - and completed reportable outputs; no unavailable component is omitted.
 - says: Report the entire source-aligned result vector, including mixed estimates and intervals. Complete coverage is not evidence of agreement, significance, equivalence, causality or successful replication.
- key: otherwise
 - region: otherwise
 - says: State the precise unresolved interpretation/coverage issue and any available component evidence without forcing it into a scientific sign or replication-success category. Unrun/operational failure remains a terminal, not this state.

protocol: |2

AUTHORITY AND SCOPE. Incorporate every applicable instruction in

- ↪ prereg/data-contract.md (schema table, ordered construction,
- ↪ inference and output contract), replication/clinical-codes.yaml and
- ↪ replication/deviations.md. These are part of this protocol, not
- ↪ optional implementation advice. P/S/R are the checked source locators
- ↪ defined in source-map.md. Only explicit source rules or lossless
- ↪ engineering may be implemented. Each STOP_UNRESOLVED_SOURCE /
- ↪ STOP_UNAVAILABLE_INPUT emits a component-specific coverage row and
- ↪ absent numerical output. A shared upstream gap propagates to all and
- ↪ only its consumers. Do not treat this stage's documented gap as a
- ↪ realized Atom result: all Atoms remain prereg and pending.

COMMON POPULATION / TIME. Women born1950-1977, alive and resident at45;
 ↪ immigration before25; no prior source contraindication, MHT or
 ↪ bilateral oophorectomy. Unit is woman's (start,stop] age interval;
 ↪ endpoint is CPR death, censoring as P3, end2023-07-31 except cause-sp
 ↪ ecific2022-12-31. Exact calendar/birthday and migration conflicts
 ↪ D01/D02, source-file mapping D03/D05 and applicable covariates
 ↪ D06/D07 must be resolved before consuming estimates. No missing key,
 ↪ date or excluded-history gap may be converted to zero/no-history. The
 ↪ source ever-use process starts unexposed and never returns to
 ↪ unexposed; source cumulative-duration process and its unresolved
 ↪ episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95%
 ↪ intervals and source software defaults in data-contract.md. P's
 ↪ two-sided .05 convention may describe individual source-specified
 ↪ tests; it does not adjudicate replication. No power calculation,
 ↪ multiplicity correction, alternative estimator, extra subgroup or
 ↪ empirical threshold is added. Use results/atom-000001/estimates.csv,
 ↪ summaries.csv, coverage.csv, adjudication.md and
 ↪ source-comparison.csv with the shared exact schemas. Do not create a
 ↪ current result file from published values. At runtime compute each
 ↪ source-defined output independently once its required inputs/rules
 ↪ are resolved; an adjusted fit's failure never overwrites
 ↪ crude/summary siblings. Retain exact target/exposure/component
 ↪ labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has termi
 ↪ nal no_result and no registered local state. For an evidence-backed
 ↪ completed module, apply verdicts in the listed first-matching order.
 ↪ Boundary/partial/complete concern coverage only. The scientific
 ↪ evidence is the entire labelled estimate/CI/descriptive record. Mixed
 ↪ signs and uncertain estimates remain mixed; missing results are not
 ↪ neutral. No local key means causal validity, significance, equivalenc
 ↪ e, replication success or independent review. Shared cohort joins and
 ↪ checks are inside this protocol, not separate experimental nodes.

SOURCE TARGETS. P Results pp4-6, Figure1, Table1, Table2 primary/duration
 ↪ /contraindication/as-treated blocks, discussion p8 follow-up and
 ↪ contraindication descriptions; S2 product distribution and S figure1;
 ↪ R279-299,429-480,485-500,525-542,632-667,698-725,806-967. Code-only
 ↪ checks are explicitly distinguished from published results. Original
 ↪ output order is target-inventory.md, not R execution order.

EXECUTABLE PSEUDOCODE.

1. REQUIRE_SOURCE for each output. Reconstruct only source-resolved
 ↪ logical inputs and shared cohort/intervals. Retain exclusion flags
 ↪ separately and union for Figure1. Emit flow.csv keyed by `node` with
 ↪ `n,status,reason` for eligible, excluded_total, thrombophilia, liver,
 ↪ arterial, venous, breast_gyn, previous_mht, previous_bso, included,
 ↪ never, ever, duration_lt1, duration_1_2_9, duration_3_4_9, duration_5
 ↪ _9_9, duration_ge10. Count eligibility before the exclusions; reasons
 ↪ overlap. Exposed and unexposed are end-of-follow-up groups for the
 ↪ flow, not baseline exposure groups in Cox.

2. Build last interval per person ordered descending tstop (R806); sum
 ↪ risktime per person and compute median,25th,75th percentiles (R817)
 ↪ in years, preserve source R quantile behavior once pinned. Emit total
 ↪ deaths/proportion, total end-of-follow-up ever-use/proportion,
 ↪ ever-use proportions for1950/1960/1970 birth cohorts, exposed
 ↪ cumulative-duration median/IQR and final duration counts/proportions,
 ↪ and end-age median in summaries.csv. End-age median P8 is a
 ↪ descriptive target using each woman's final tstop, not a survival
 ↪ estimate. Death age is not life expectancy. If D01-D04 block their
 ↪ inputs, emit precise absent rows; no published counts enter results.
3. Landmark Table1: retain source intervals starting exactly55 (R950),
 ↪ group by past/present MHT at55. Assert one valid landmark row per
 ↪ included woman; otherwise stop Table1 rather than count duplicates.
 ↪ Report the exact characteristic roster in target-catalog.yaml:
 ↪ country including unknown, parity mean/SD and0/1-2/>=3, marriage
 ↪ status incl unknown, four education levels+unknown, imputed income
 ↪ q1-q4, calendar periods2002-08/2009-15/2016-23, six individually
 ↪ listed cardiometabolic indicators, >=3 contacts, prior hyst/BS0. Use
 ↪ characteristics.csv:
 ↪ `characteristic,category,exposure,n,percent,mean,sd,status,reason`;
 ↪ no t-tests (R953 test=F). Denominators are group-specific landmark N,
 ↪ unknowns explicit except source-disclosure income imputation.
 ↪ Missingness proportions from each woman's final row R963-967 are
 ↪ separate scalar targets, not Table1 missingness estimates.
4. At55 calculate proportions for [2004,2007) and[2021,2024) as source
 ↪ R957-960, not all2007-2020 as a new trend. S2 requested cells are
 ↪ products.csv: `atc,regimen,form,percent,status,reason`, fixed
 ↪ source29-row roster. Source R430 groups nonmissing htatc rows'
 ↪ risktime and divides by total; P/S describe proportion of MHT. D19
 ↪ stops the proportion target until the denominator is established; do
 ↪ not replace with dose or prescription counts.
5. Population main binary: summary by ht=0/1, events=sum(event),
 ↪ person_years=sum(risktime), IR=10000*events/person_years. Crude m01:
 ↪ Surv(tstart,tstop,event)~ht. Adjusted m02 R461: ~ht+year_cat + fland_
 ↪ mode + udda_mode + q_ind_mode + parity_cat + dm + hypertens + afli +
 ↪ valv + hf + statin + many_diag + civst_mode. Printed adjustment list
 ↪ is A_P=year_cat + fland_mode + udda_mode + q_ind_mode + parity_cat +
 ↪ dm + hypertens + afli + valv + hf + statin + many_diag. D08 prevents
 ↪ selecting or estimating an alternative for the published adjusted
 ↪ target. Retain crude/count siblings if their cohort/exposure inputs
 ↪ are fully resolved. No extra code-versus-paper fit is created.
6. Duration: 0,<1,1-2.9,3-4.9,5-9.9,>=10 years, with the R cut bounds
 ↪ specified in shared construction. m11 ~ht_cat; m12 ~ht_cat+A_R.
 ↪ Reference0, each nonreference coefficient/95%CI, no trend test.
 ↪ D04/epsilon/upper-bound conflicts stop affected duration assignments;
 ↪ D08 separately stops adjusted attribution.
7. Contraindication sensitivity R486: retain pre-contraindication
 ↪ intervals where thromb=vte=cvd=breastgyn=liver=0, thus censor at
 ↪ first qualifying contraindication. Summary/m21 ~ht; m22 ~ht+A_R.
 ↪ Source time-dependent boundary-day convention must match resolved
 ↪ tmerge semantics. Emit source R489 percentage of all deaths with
 ↪ preceding contraindication as separate descriptive target. Main risk
 ↪ set is not censored for these conditions.

8. As-treated R136/R526: source cessation recipe extends episode end
 - by90/365.2; compare to next episode start; mark gaps when
 - htstop90<next start, also terminal episode, select earliest marked
 - stop per person, round3 decimals. Before first MHT htstop90 is
 - forced0; retain intervals htstop90==0. Summary/m41 ~ht; m42 ~ht+A.R.
 - No reentry, no additional grace period, no drug-specific deviation.
 - D04 episode and open-end ordering must be resolved; never use
 - cessation sensitivity to replace the blocked ever-use target.
9. PH: S figure1 is14 term panels in order
 - ht,year_cat,fland_mode,civst_mode,udda_mode,q_ind_mode,parity_cat,dm,
 - hypertens,aqli,valu,hf,statin,many_diag (visually checked). R465-466
 - uses cox.zph(m02), ggcoxzph. R480 computes cox.zph(m12) without a
 - published plot; retain diagnostic existence in audit, do not add
 - another manuscript figure. After D08/D18 resolved export approved
 - `schoenfeld\_points.csv(term,time,residual)`,
 - `schoenfeld\_curves.csv(term,time,fitted,lower,upper)`,
 - `schoenfeld\_panels.csv(term,status,reason,p\_value)` and
 - `schoenfeld\_ticks.csv(term,position,label)` preserving the original
 - transformed coordinates and age labels. All scientific
 - transforms/smoothing/intervals come from the pinned original R
 - plotting method; Python draws them only. Exact ggcoxzph/survminer
 - implementation/version is an input prerequisite, not freedom to
 - choose a smoother. Report qualitative assessment and source P-values
 - without inventing PH accept/reject rules or fitting a repaired model.
10. Code-only sensitivity rows (not published Table2 values): R633/647
 - interval-start split <2003 versus>=2003; m111/m121 ~ht, m112/m122
 - ~ht+(A_R without year_cat). D09 stops unresolved calendar risk sets;
 - no2003 splitting repair. R663 m132 ~ht+A_R+hyst+bso, shares original
 - population crude/summary; source code explicitly supports these
 - adjustments only for this separate specification. R699 BMI
 - observed-interval subset, categories BMI<30 versus>=30, m161 ~ht,
 - m162 ~ht+A_R+bmi_cat. BMI measurement source/timing D06 must resolve;
 - never restrict to newly chosen measurements. Smoking R714/715 has
 - separately supported smoking-complete summary, but m171/m172/rt172
 - conflict D13: no smoking HR or merged table estimate, no repaired
 - fit. Emit exact `code\_\*` rows in a source-code appendix interface,
 - not as published numerical comparisons. All applicable shared source
 - conflicts remain gates.
11. Assert component identities and comparator labels, round only
 - presentation, apply source disclosure treatment (event counts<3
 - suppressed, not imputed). Export rows for every inventory cell.
 - Report any combination of crude/adjusted/count/PH availability; the
 - primary association remains unestimated if its specified model is
 - unresolved even when sensitivities produce results.

SCIENTIFIC MEANING. Population associations depend on

- measured-confounding control and informative-treatment/censoring
- limitations. No causal benefit, clinical safety, equivalence or
- entire-population agreement follows from precision, an HR below1 or
- completed outputs. Source-only sensitivity results are supporting
- evidence within this question, never a replacement primary endpoint.

Registered analysis module

This is the paper's central mortality question. Cohort flow, landmark characteristics and utilization descriptions orient it; duration, contraindication/cessation sensitivities, PH assessment and source-code-only calendar/BMI/surgery-adjusted checks interrogate that same association. They are not separate experiments.

All empirical results and local registered state are unfilled. The numerical design ratings are native serialization compatibility fields, not scientific evidence, feasibility validation, novelty competition or a replication score. Feasibility1 reflects explicit unresolved source/input dependencies; no independent scientific acceptance is claimed.

Reference implementation

Entry point and supported/unresolved components for `atom-000001`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000002: Sibling mortality

AUTHORITY AND SCOPE. Incorporate every applicable instruction in `prereg/data-contract.md` (schema table, ordered construction, inference and output contract), `replication/clinical-codes.yaml` and `replication/deviations.md`. These are part of this protocol, not optional implementation advice. P/S/R are the checked source locators defined in `source-map.md`. Only explicit source rules or loss-less engineering may be implemented. Each `STOP_UNRESOLVED_SOURCE` / `STOP_UNAVAILABLE_INPUT` emits a component-specific coverage row and absent numerical output. A shared upstream gap propagates to all and only its consumers. Do not treat this stage's documented gap as a realized Atom result: all Atoms remain prereg and pending.

COMMON POPULATION / TIME. Women born1950–1977, alive and resident at45; immigration before25; no prior source contraindication, MHT or bilateral oophorectomy. Unit is woman's (start,stop] age interval; endpoint is CPR death, censoring as P3, end2023-07-31 except cause-specific2022-12-31. Exact calendar/birthday and migration conflicts D01/D02, source-file mapping D03/D05 and applicable covariates D06/D07 must be resolved before consuming estimates. No missing key, date or excluded-history gap may be converted to zero/no-history. The source ever-use process starts unexposed and never returns to unexposed; source cumulative-duration process and its unresolved episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95% intervals and source software defaults in `data-contract.md`. P's two-sided .05 convention may describe individual source-specified tests; it does not adjudicate replication. No power calculation, multiplicity correction, alternative estimator, extra subgroup or empirical threshold is added. Use `results/atom-000002/estimates.csv`, `summaries.csv`, `coverage.csv`, `adjudication.md` and `source-comparison.csv` with the shared exact schemas. Do not create a current result file from published values. At runtime compute each source-defined output independently once its required inputs/rules are resolved; an adjusted fit's failure never overwrites crude/summary siblings. Retain exact target/exposure/component labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has terminal `no_result` and no registered local state. For an evidence-backed completed module, apply verdicts in the listed first-matching order. Boundary/partial/complete concern

coverage only. The scientific evidence is the entire labelled estimate/CI/descriptive record. Mixed signs and uncertain estimates remain mixed; missing results are not neutral. No local key means causal validity, significance, equivalence, replication success or independent review. Shared cohort joins and checks are inside this protocol, not separate experimental nodes.

SOURCE. P3 sibling design, P4/Table2 sibling result, R43–47,502–523,829–830.

EXECUTABLE PSEUDOCODE. 1. REQUIRE_SOURCE(shared cohort, binary exposure, full-person mother links/cleaning). From final interval per eligible woman, ht_sib=1 if ht_sum>0 else0. Join source parent identifiers. Require nonmissing mother; count sisters by mother>1; retain mother groups with max(ht_sib)=1 and max(1-ht_sib)=1. Shared father is not required. R's parent>30 rule is evaluated in the original imported person universe; no substitute count in the reduced cohort. 2. Assign family_id to the common mother, retain all eligible sisters in each retained family, with each woman's entire original time-varying history. Join retained person/family back many-intervals-to-one-woman, assert no multiplication. Discordance at end is source-defined selection using future exposure; do not recast as baseline randomization or incident sibling matching. Summaries: total women and percentage of main cohort, events/person-years/IR by contemporaneous ht. Later-exposed women contribute unexposed time until first use. 3. Source crude m31 R515 ~ht in sib2 does not include strata. P says the comparison uses mother strata. D11 blocks the printed crude-within-family target's interpretation/model resolution. Do not add strata or call this unstratified source model within-family. Counts/person-years remain independently reportable if prerequisites are resolved. 4. Adjusted source m32 R518: Surv(tstart,tstop,event)~ht+year_cat + fland_mode + udda_mode + q_ind_mode + parity_cat + dm + hypertens + affi + valv + hf + statin + many_diag + civst_mode+strata(family_id). Source uses mother-specific baseline hazards, not a frailty model, sister pair difference, robust cluster variance or family fixed-effect coefficient. D08 adjustment-set attribution plus D06/D07 covariates stop this printed adjusted target until clarified. Retain exact requested model extraction and absent component row; do not run alternative adjustment sets. 5. Export target_id=t2_sibling estimates rows for never/past-present and summary/crude/adjusted; sibling_population summaries with n and share_pct; coverage/adjudication. If links are unavailable emit boundary without claiming the entire nationwide study failed. Fixed TRUE activation means its execution is not selected using main HR; shared input failure remains an explicit gate inside this protocol.

SCIENTIFIC MEANING. Compare the family-specific adjusted HR with the population HR only as separate associations. Shared-family confounding may be reduced; nonshared confounding, selection by eventual discordance, time-varying health and treatment indication remain. A wide CI is imprecise evidence, not successful confirmation, and this sample is not interchangeable with the nationwide cohort. No formal between-model difference test or equivalence band is added.

Complete registration fields

```
id: atom-000002
topic: Sibling mortality
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds:
- target: atom-000001
  bond: prereg
  description: Planned interpretation relationship to population mortality;
    not an input dependency or outcome-dependent activation.
preregistration:
  version: 1
  mode: canonical
  question: What mortality association is identified by the source comparison
    of sisters discordant in eventual MHT use?
  rationale: The sibling comparison changes the identifying comparison to
    families selected for discordant exposure. Its judgment about shared
    ↪ familial
    confounding is interpretable independently from the nationwide
    ↪ association;
    it warrants its own Atom, while its crude/adjusted variants and
    ↪ descriptive
    size remain together.
  assumption: The source question is observational. Correct record linkage,
    source-equivalent histories, treatment/censoring specification and
    ↪ reported-model
    attribution are prerequisites; neither schema nor source agreement proves
    absence of confounding.
  target_strength: associational
  target_shape: association
  ratings:
    importance: 3
    novelty: 1
    feasibility: 1
    data_advantage: 3
    informativeness: 0.5
  tables:
    - BEF
    - VNDS
    - DOD
    - LMDB
    - LPR_ADM
    - LPR_DIAG
    - LPR_OPR
    - LPR_SKSOPR
    - LPR_A_KONTAKT
    - LPR_A_DIAGNOSE
    - LPR_A_PROCREGISTRERING
    - MFR
    - CIV
    - IND
    - UDDA
    - UDDF
  activation: 'TRUE'
  verdicts:
    - key: boundary
```

region: After a completed evidence-backed assessment, no requested
 ↳ empirical
 component can be produced because every component is
 ↳ source-underdetermined
 or unavailable. Operational failure is excluded from this region.
 says: Report precise no-estimate boundaries. There is no empirical
 ↳ mortality
 finding or replication-agreement conclusion.

- key: partial
 region: At least one requested empirical component is available and at
 least one is explicitly unresolved, unavailable, not estimable or
 ↳ disclosure-suppressed
 after completed assessment. No operationally unfinished component is
 hidden by this state.
 says: Report the available target-specific estimates, intervals and
 ↳ descriptions
 and all unavailable components. Partial coverage is not statistical
 confirmation or a uniform evidence direction.

- key: complete
 region: All requested empirical components within this Atom, including
 code-only targets explicitly assigned to it, have resolved source
 ↳ specifications
 and completed reportable outputs; no unavailable component is omitted.
 says: Report the entire source-aligned result vector, including mixed
 estimates and intervals. Complete coverage is not evidence of agreement,
 significance, equivalence, causality or successful replication.

- key: otherwise
 region: otherwise
 says: State the precise unresolved interpretation/coverage issue and any
 available component evidence without forcing it into a scientific sign
 or replication-success category. Unrun/operational failure remains a
 terminal, not this state.

protocol: |2

AUTHORITY AND SCOPE. Incorporate every applicable instruction in
 ↳ prereg/data-contract.md (schema table, ordered construction,
 ↳ inference and output contract), replication/clinical-codes.yaml and
 ↳ replication/deviations.md. These are part of this protocol, not
 ↳ optional implementation advice. P/S/R are the checked source locators
 ↳ defined in source-map.md. Only explicit source rules or lossless
 ↳ engineering may be implemented. Each STOP_UNRESOLVED_SOURCE /
 ↳ STOP_UNAVAILABLE_INPUT emits a component-specific coverage row and
 ↳ absent numerical output. A shared upstream gap propagates to all and
 ↳ only its consumers. Do not treat this stage's documented gap as a
 ↳ realized Atom result: all Atoms remain prereg and pending.

COMMON POPULATION / TIME. Women born1950-1977, alive and resident at45;
 ↳ immigration before25; no prior source contraindication, MHT or
 ↳ bilateral oophorectomy. Unit is woman's (start,stop] age interval;
 ↳ endpoint is CPR death, censoring as P3, end2023-07-31 except cause-sp
 ↳ ecific2022-12-31. Exact calendar/birthday and migration conflicts
 ↳ D01/D02, source-file mapping D03/D05 and applicable covariates
 ↳ D06/D07 must be resolved before consuming estimates. No missing key,
 ↳ date or excluded-history gap may be converted to zero/no-history. The
 ↳ source ever-use process starts unexposed and never returns to
 ↳ unexposed; source cumulative-duration process and its unresolved
 ↳ episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95%
 ↪ intervals and source software defaults in data-contract.md. P's
 ↪ two-sided .05 convention may describe individual source-specified
 ↪ tests; it does not adjudicate replication. No power calculation,
 ↪ multiplicity correction, alternative estimator, extra subgroup or
 ↪ empirical threshold is added. Use results/atom-000002/estimates.csv,
 ↪ summaries.csv, coverage.csv, adjudication.md and
 ↪ source-comparison.csv with the shared exact schemas. Do not create a
 ↪ current result file from published values. At runtime compute each
 ↪ source-defined output independently once its required inputs/rules
 ↪ are resolved; an adjusted fit's failure never overwrites
 ↪ crude/summary siblings. Retain exact target/exposure/component
 ↪ labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has termi
 ↪ nal no_result and no registered local state. For an evidence-backed
 ↪ completed module, apply verdicts in the listed first-matching order.
 ↪ Boundary/partial/complete concern coverage only. The scientific
 ↪ evidence is the entire labelled estimate/CI/descriptive record. Mixed
 ↪ signs and uncertain estimates remain mixed; missing results are not
 ↪ neutral. No local key means causal validity, significance, equivalenc
 ↪ e, replication success or independent review. Shared cohort joins and
 ↪ checks are inside this protocol, not separate experimental nodes.

SOURCE. P3 sibling design, P4/Table2 sibling result,
 ↪ R43-47,502-523,829-830.

EXECUTABLE PSEUDOCODE.

1. REQUIRE_SOURCE(shared cohort, binary exposure, full-person mother
 ↪ links/cleaning). From final interval per eligible woman, ht_sib=1 if
 ↪ ht_sum>0 else0. Join source parent identifiers. Require nonmissing
 ↪ mother; count sisters by mother>1; retain mother groups with
 ↪ max(ht_sib)=1 and max(1-ht_sib)=1. Shared father is not required. R's
 ↪ parent>30 rule is evaluated in the original imported person universe;
 ↪ no substitute count in the reduced cohort.
2. Assign family_id to the common mother, retain all eligible sisters in
 ↪ each retained family, with each woman's entire original time-varying
 ↪ history. Join retained person/family back many-intervals-to-one-woman
 ↪ , assert no multiplication. Discordance at end is source-defined
 ↪ selection using future exposure; do not recast as baseline randomizat
 ↪ ion or incident sibling matching. Summaries: total women and
 ↪ percentage of main cohort, events/person-years/IR by contemporaneous
 ↪ ht. Later-exposed women contribute unexposed time until first use.
3. Source crude m31 R515 ~ht in sib2 does not include strata. P says the
 ↪ comparison uses mother strata. D11 blocks the printed
 ↪ crude-within-family target's interpretation/model resolution. Do not
 ↪ add strata or call this unstratified source model within-family.
 ↪ Counts/person-years remain independently reportable if prerequisites
 ↪ are resolved.
4. Adjusted source m32 R518: Surv(tstart,tstop,event)~ht+year_cat +
 ↪ fland_mode + udda_mode + q_ind_mode + parity_cat + dm + hypertens +
 ↪ afli + valv + hf + statin + many_diag + civst_mode+strata(family_id).
 ↪ Source uses mother-specific baseline hazards, not a frailty model,
 ↪ sister pair difference, robust cluster variance or family
 ↪ fixed-effect coefficient. D08 adjustment-set attribution plus D06/D07
 ↪ covariates stop this printed adjusted target until clarified. Retain
 ↪ exact requested model extraction and absent component row; do not run
 ↪ alternative adjustment sets.

```

5. Export `target_id=t2_sibling` estimates rows for never/past-present
→ and summary/crude/adjusted; `sibling_population` summaries with n and
→ share_pct; coverage/adjudication. If links are unavailable emit
→ boundary without claiming the entire nationwide study failed. Fixed
→ TRUE activation means its execution is not selected using main HR;
→ shared input failure remains an explicit gate inside this protocol.

```

```

SCIENTIFIC MEANING. Compare the family-specific adjusted HR with the
→ population HR only as separate associations. Shared-family
→ confounding may be reduced; nonshared confounding, selection by
→ eventual discordance, time-varying health and treatment indication
→ remain. A wide CI is imprecise evidence, not successful confirmation,
→ and this sample is not interchangeable with the nationwide cohort. No
→ formal between-model difference test or equivalence band is added.

```

Registered analysis module

The sibling comparison changes the identifying comparison to families selected for discordant exposure. Its judgment about shared familial confounding is interpretable independently from the nationwide association; it warrants its own Atom, while its crude/adjusted variants and descriptive size remain together.

All empirical results and local registered state are unfilled. The numerical design ratings are native serialization compatibility fields, not scientific evidence, feasibility validation, novelty competition or a replication score. Feasibility1 reflects explicit unresolved source/input dependencies; no independent scientific acceptance is claimed.

Reference implementation

Entry point and supported/unresolved components for atom-000002: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000003: Surgical subgroup mortality

AUTHORITY AND SCOPE. Incorporate every applicable instruction in prereg/data-contract.md (schema table, ordered construction, inference and output contract), replication/clinical-codes.yaml and replication/deviations.md. These are part of this protocol, not optional implementation advice. P/S/R are the checked source locators defined in source-map.md. Only explicit source rules or loss-less engineering may be implemented. Each STOP_UNRESOLVED_SOURCE / STOP_UNAVAILABLE_INPUT emits a component-specific coverage row and absent numerical output. A shared upstream gap propagates to all and only its consumers. Do not treat this stage's documented gap as a realized Atom result: all Atoms remain prereg and pending.

COMMON POPULATION / TIME. Women born1950–1977, alive and resident at45; immigration before25; no prior source contraindication, MHT or bilateral oophorectomy. Unit is woman's (start,stop] age interval; endpoint is CPR death, censoring as P3, end2023-07-31 except cause-specific2022-12-31. Exact calendar/birthday

and migration conflicts D01/D02, source-file mapping D03/D05 and applicable covariates D06/D07 must be resolved before consuming estimates. No missing key, date or excluded-history gap may be converted to zero/no-history. The source ever-use process starts unexposed and never returns to unexposed; source cumulative-duration process and its unresolved episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95% intervals and source software defaults in data-contract.md. P's two-sided .05 convention may describe individual source-specified tests; it does not adjudicate replication. No power calculation, multiplicity correction, alternative estimator, extra subgroup or empirical threshold is added. Use results/atom-000003/estimates.csv, summaries.csv, coverage.csv, adjudication.md and source-comparison.csv with the shared exact schemas. Do not create a current result file from published values. At runtime compute each source-defined output independently once its required inputs/rules are resolved; an adjusted fit's failure never overwrites crude/summary siblings. Retain exact target/exposure/component labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has terminal no_result and no registered local state. For an evidence-backed completed module, apply verdicts in the listed first-matching order. Boundary/partial/complete concern coverage only. The scientific evidence is the entire labelled estimate/CI/descriptive record. Mixed signs and uncertain estimates remain mixed; missing results are not neutral. No local key means causal validity, significance, equivalence, replication success or independent review. Shared cohort joins and checks are inside this protocol, not separate experimental nodes.

SOURCE. P3 surgery methods; P4-6 subgroup and death-age results; Table2 surgery blocks; R568-613,669-695. See D05/D12 for event reconstruction and risk-set conflicts.

EXECUTABLE PSEUDOCODE. 1. REQUIRE_SOURCE(cohort, surgery events/dates, exposure duration, breast/gynaecological cancer onset). Hysterectomy and BSO need not be mutually exclusive. Carry source indicators with tmerge; censor at breast/endometrial/ovarian cancer according to supplied C50/C54-57 code set and resolved event timing. Other contraindications do not censor these groups unless present in the original common cohort exclusion. Do not add postoperative washout or new eligibility criteria. 2. Hyst-any R569: keep breastgyn==0 and hyst==1 intervals; use 0/<5/>=5 duration categories as source shared cuts. Summaries by category, m71 ~ht_cat2, m72 ~ht_cat2+year_cat + fland_mode + udda_mode + q_ind_mode + parity_cat + dm + hypertens + affi + valv + hf + statin + many_diag + civst_mode. Women can enter surgical risk sets when hyst indicator turns1, including at entry for previous hysterectomy; no additional restriction by MHT start relative to surgery. D08 stops adjusted attribution, not independently valid counts/crude. 3. BSO45-54: R583 starts from time-dependent bso, overwrites bso2 for eventual>=55. For women with source bso_age<55, bso==1 identifies postoperative intervals after eligible BSO at>=45 (baseline exclusions apply). Retain breastgyn==0; summaries, m81 ~ht_cat2, m82 ~ht_cat2+A_R. Source D01 at45 and D05 second-unilateral

ambiguity gate affected risk sets. Do not estimate a new treatment-initiation-at-surgery trial. 4. BSO $\geq$ 55: source R583 assigns bso=2 for all history of eventual older-surgery women, including before operation. P's had-undergone/time-varying specification does not establish that retrospective risk set. D12: emit no-estimate summary/crude/adjusted rows for t2_bso_ge55 until source evidence resolves timing; neither literal preoperative inclusion nor repaired postoperative filtering is authorized. The source m91/m92 formula is $\sim\text{ht_cat2} / \sim\text{ht_cat2} + \text{A_R}$; retain it as extraction, not permission to run it on a guessed risk set. 5. Death-age target R597-599: in the source BSO45-54, breastgyn-censored risk set, count event==1 and report tstop median/25th/75th percentile separately by ht at death, corresponding to P5's subsequently-deceased women. Emit **bso_death_age** summaries for user/nonuser in years and **bso_deaths** count. This is conditional on dying, not mean survival/life-years gained. Do not assign the published703 count to the surgical population. 6. Code-only hyst strata R670: extract hysterectomy-age<55 versus $\geq$ 55 analyses, collapsed duration, breastgyn censoring; m141/m151 crude and m142/m152 +A_R. The comment45-54 does not exclude previous hysterectomy; code hyst==1 may include it. D12 blocks intended age-specific outputs until population/time origin clarified; dropped bso sorting is only display defect D16. No guessed45-year restriction. Keep these requested source-code-only targets visible, without inventing published references. 7. Export estimates for t2_hyst, t2_bso_45_54, t2_bso_ge55, code_hyst_lt55, code_hyst_ge55, separate groups0/<5/>=5 and components; summaries and coverage. Retain partial scope even if older group fails. No pooled surgery coefficient or test between overlapping populations.

SCIENTIFIC MEANING. Jointly describe the surgery-specific associations and their differences in scope, precision and availability. Indications for surgery, treatment selection and cancer censoring constrain interpretation. Neither a younger-surgery association nor a death-age contrast resolves the older group's source conflict or demonstrates causal mortality reduction.

Complete registration fields

```
id: atom-000003
topic: Surgical subgroup mortality
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds:
- target: atom-000001
  bond: prereg
  description: Planned interpretation relationship to population mortality;
    not an input dependency or outcome-dependent activation.
preregistration:
  version: 1
  mode: canonical
  question: How does the source mortality association differ in its
    ↪ hysterectomy
    and bilateral-oophorectomy populations?
  rationale: The surgery analyses address a clinically distinct population
```

and timing of ovarian removal. Hysterectomy, BS045-54 and BS0>=55 are related contrasts within that question, with overlapping memberships and separately preserved risk sets/results. Age-stratified hysterectomy code variants remain in this module; an unsupported older group cannot erase supported younger-group evidence.

assumption: The source question is observational. Correct record linkage, source-equivalent histories, treatment/censoring specification and
↳ reported-model attribution are prerequisites; neither schema nor source agreement proves absence of confounding.

target_strength: associational
target_shape: association

ratings:
importance: 3
novelty: 1
feasibility: 1
data_advantage: 3
informativeness: 0.5

tables:
- BEF
- VNDS
- DOD
- LMDB
- LPR_ADM
- LPR_DIAG
- LPR_OPR
- LPR_SKSOPR
- LPR_A_KONTAKT
- LPR_A_DIAGNOSE
- LPR_A_PROCREGISTRERING
- MFR
- CIV
- IND
- UDDA
- UDDF

activation: 'TRUE'

verdicts:
- key: boundary
region: After a completed evidence-backed assessment, no requested
↳ empirical component can be produced because every component is
↳ source-underdetermined or unavailable. Operational failure is excluded from this region.
says: Report precise no-estimate boundaries. There is no empirical
↳ mortality finding or replication-agreement conclusion.

- key: partial
region: At least one requested empirical component is available and at least one is explicitly unresolved, unavailable, not estimable or
↳ disclosure-suppressed after completed assessment. No operationally unfinished component is hidden by this state.
says: Report the available target-specific estimates, intervals and
↳ descriptions and all unavailable components. Partial coverage is not statistical confirmation or a uniform evidence direction.

- key: complete
region: All requested empirical components within this Atom, including code-only targets explicitly assigned to it, have resolved source
↳ specifications

and completed reportable outputs; no unavailable component is omitted.
 says: Report the entire source-aligned result vector, including mixed estimates and intervals. Complete coverage is not evidence of agreement, significance, equivalence, causality or successful replication.
 - key: otherwise
 region: otherwise
 says: State the precise unresolved interpretation/coverage issue and any available component evidence without forcing it into a scientific sign or replication-success category. Unrun/operational failure remains a terminal, not this state.
 protocol: |2

AUTHORITY AND SCOPE. Incorporate every applicable instruction in
 → prereg/data-contract.md (schema table, ordered construction,
 → inference and output contract), replication/clinical-codes.yaml and
 → replication/deviations.md. These are part of this protocol, not
 → optional implementation advice. P/S/R are the checked source locators
 → defined in source-map.md. Only explicit source rules or lossless
 → engineering may be implemented. Each STOP_UNRESOLVED_SOURCE /
 → STOP_UNAVAILABLE_INPUT emits a component-specific coverage row and
 → absent numerical output. A shared upstream gap propagates to all and
 → only its consumers. Do not treat this stage's documented gap as a
 → realized Atom result: all Atoms remain prereg and pending.

COMMON POPULATION / TIME. Women born1950-1977, alive and resident at45;
 → immigration before25; no prior source contraindication, MHT or
 → bilateral oophorectomy. Unit is woman's (start,stop] age interval;
 → endpoint is CPR death, censoring as P3, end2023-07-31 except cause-sp
 → ecific2022-12-31. Exact calendar/birthday and migration conflicts
 → D01/D02, source-file mapping D03/D05 and applicable covariates
 → D06/D07 must be resolved before consuming estimates. No missing key,
 → date or excluded-history gap may be converted to zero/no-history. The
 → source ever-use process starts unexposed and never returns to
 → unexposed; source cumulative-duration process and its unresolved
 → episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95%
 → intervals and source software defaults in data-contract.md. P's
 → two-sided .05 convention may describe individual source-specified
 → tests; it does not adjudicate replication. No power calculation,
 → multiplicity correction, alternative estimator, extra subgroup or
 → empirical threshold is added. Use results/atom-000003/estimates.csv,
 → summaries.csv, coverage.csv, adjudication.md and
 → source-comparison.csv with the shared exact schemas. Do not create a
 → current result file from published values. At runtime compute each
 → source-defined output independently once its required inputs/rules
 → are resolved; an adjusted fit's failure never overwrites
 → crude/summary siblings. Retain exact target/exposure/component
 → labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has termi
 → nal no_result and no registered local state. For an evidence-backed
 → completed module, apply verdicts in the listed first-matching order.
 → Boundary/partial/complete concern coverage only. The scientific
 → evidence is the entire labelled estimate/CI/descriptive record. Mixed
 → signs and uncertain estimates remain mixed; missing results are not
 → neutral. No local key means causal validity, significance, equivalenc
 → e, replication success or independent review. Shared cohort joins and
 → checks are inside this protocol, not separate experimental nodes.

SOURCE. P3 surgery methods; P4-6 subgroup and death-age results; Table2
 ↳ surgery blocks; R568-613,669-695. See D05/D12 for event
 ↳ reconstruction and risk-set conflicts.

EXECUTABLE PSEUDOCODE.

1. REQUIRE_SOURCE(cohort, surgery events/dates, exposure duration,
 ↳ breast/gynaecological cancer onset). Hysterectomy and BSO need not be
 ↳ mutually exclusive. Carry source indicators with tmerge; censor at br
 ↳ east/endometrial/ovarian cancer according to supplied C50/C54-57 code
 ↳ set and resolved event timing. Other contraindications do not censor
 ↳ these groups unless present in the original common cohort exclusion.
 ↳ Do not add postoperative washout or new eligibility criteria.
2. Hyst-any R569: keep breastgyn==0 and hyst==1 intervals; use 0/<5/>=5
 ↳ duration categories as source shared cuts. Summaries by category, m71
 ↳ ~ht_cat2, m72 ~ht_cat2+year_cat + fland_mode + udda_mode + q_ind_mode
 ↳ + parity_cat + dm + hypertens + afli + valv + hf + statin + many_diag
 ↳ + civst_mode. Women can enter surgical risk sets when hyst indicator
 ↳ turns1, including at entry for previous hysterectomy; no additional
 ↳ restriction by MHT start relative to surgery. D08 stops adjusted
 ↳ attribution, not independently valid counts/crude.
3. BS045-54: R583 starts from time-dependent bso, overwrites bso2 for
 ↳ eventual>=55. For women with source bso_age<55, bso==1 identifies
 ↳ postoperative intervals after eligible BSO at>=45 (baseline
 ↳ exclusions apply). Retain breastgyn==0; summaries, m81 ~ht_cat2, m82
 ↳ ~ht_cat2+A_R. Source D01 at45 and D05 second-unilateral ambiguity
 ↳ gate affected risk sets. Do not estimate a new
 ↳ treatment-initiation-at-surgery trial.
4. BS0>=55: source R583 assigns bso=2 for all history of eventual
 ↳ older-surgery women, including before operation. P's had-undergone/ti
 ↳ me-varying specification does not establish that retrospective risk
 ↳ set. D12: emit no-estimate summary/crude/adjusted rows for t2_bso_ge5
 ↳ 5 until source evidence resolves timing; neither literal preoperative
 ↳ inclusion nor repaired postoperative filtering is authorized. The
 ↳ source m91/m92 formula is ~ht_cat2 / ~ht_cat2+A_R; retain it as
 ↳ extraction, not permission to run it on a guessed risk set.
5. Death-age target R597-599: in the source BS045-54, breastgyn-censored
 ↳ risk set, count event==1 and report tstop median/25th/75th percentile
 ↳ separately by ht at death, corresponding to P5's
 ↳ subsequently-deceased women. Emit `bso\_death\_age` summaries for
 ↳ user/nonuser in years and `bso\_deaths` count. This is conditional on
 ↳ dying, not mean survival/life-years gained. Do not assign the
 ↳ published703 count to the surgical population.
6. Code-only hyst strata R670: extract hysterctomy-age<55 versus>=55
 ↳ analyses, collapsed duration, breastgyn censoring; m141/m151 crude
 ↳ and m142/m152 +A_R. The comment45-54 does not exclude previous
 ↳ hysterectomy; code hyst==1 may include it. D12 blocks intended
 ↳ age-specific outputs until population/time origin clarified; dropped
 ↳ bso sorting is only display defect D16. No guessed45-year
 ↳ restriction. Keep these requested source-code-only targets visible,
 ↳ without inventing published references.
7. Export estimates for `t2\_hyst`, `t2\_bso\_45\_54`, `t2\_bso\_ge55`,
 ↳ `code\_hyst\_lt55`, `code\_hyst\_ge55`, separate groups0/<5/>=5 and
 ↳ components; summaries and coverage. Retain partial scope even if
 ↳ older group fails. No pooled surgery coefficient or test between
 ↳ overlapping populations.

SCIENTIFIC MEANING. Jointly describe the surgery-specific associations
 ↳ and their differences in scope, precision and availability.
 ↳ Indications for surgery, treatment selection and cancer censoring
 ↳ constrain interpretation. Neither a younger-surgery association nor a
 ↳ death-age contrast resolves the older group's source conflict or
 ↳ demonstrates causal mortality reduction.

Registered analysis module

The surgery analyses address a clinically distinct population and timing of ovarian removal. Hysterectomy, BSO45–54 and BSO≥55 are related contrasts within that question, with overlapping memberships and separately preserved risk sets/results. Age-stratified hysterectomy code variants remain in this module; an unsupported older group cannot erase supported younger-group evidence.

All empirical results and local registered state are unfilled. The numerical design ratings are native serialization compatibility fields, not scientific evidence, feasibility validation, novelty competition or a replication score. Feasibility1 reflects explicit unresolved source/input dependencies; no independent scientific acceptance is claimed.

Reference implementation

Entry point and supported/unresolved components for atom-000003: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000004: Treatment heterogeneity

AUTHORITY AND SCOPE. Incorporate every applicable instruction in prereg/data-contract.md (schema table, ordered construction, inference and output contract), replication/clinical-codes.yaml and replication/deviations.md. These are part of this protocol, not optional implementation advice. P/S/R are the checked source locators defined in source-map.md. Only explicit source rules or loss-less engineering may be implemented. Each STOP_UNRESOLVED_SOURCE / STOP_UNAVAILABLE_INPUT emits a component-specific coverage row and absent numerical output. A shared upstream gap propagates to all and only its consumers. Do not treat this stage's documented gap as a realized Atom result: all Atoms remain prereg and pending.

COMMON POPULATION / TIME. Women born1950–1977, alive and resident at45; immigration before25; no prior source contraindication, MHT or bilateral oophorectomy. Unit is woman's (start,stop] age interval; endpoint is CPR death, censoring as P3, end2023-07-31 except cause-specific2022-12-31. Exact calendar/birthday and migration conflicts D01/D02, source-file mapping D03/D05 and applicable covariates D06/D07 must be resolved before consuming estimates. No missing key, date or excluded-history gap may be converted to zero/no-history. The source ever-use process starts unexposed and never returns to unexposed; source cumulative-duration process and its unresolved episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95% intervals and source software defaults in data-contract.md. P's two-sided .05 convention may describe individual source-specified tests; it does not adjudicate replication. No power calculation, multiplicity correction, alternative estimator, extra subgroup or empirical threshold is added. Use results/atom-000004/estimates.csv, summaries.csv, coverage.csv, adjudication.md and source-comparison.csv with the shared exact schemas. Do not create a current result file from published values. At runtime compute each source-defined output independently once its required inputs/rules are resolved; an adjusted fit's failure never overwrites crude/summary siblings. Retain exact target/exposure/component labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has terminal no_result and no registered local state. For an evidence-backed completed module, apply verdicts in the listed first-matching order. Boundary/partial/complete concern coverage only. The scientific evidence is the entire labelled estimate/CI/descriptive record. Mixed signs and uncertain estimates remain mixed; missing results are not neutral. No local key means causal validity, significance, equivalence, replication success or independent review. Shared cohort joins and checks are inside this protocol, not separate experimental nodes.

SOURCE. P3-5 stratified analyses, Table2 final three blocks; R384-427,544-566,615-630.

EXECUTABLE PSEUDOCODE. 1. REQUIRE_SOURCE(shared eligible history, dose units and episode assignment). Before MHT use, all categories reference never-used. Keep total cumulative MHT and cumulative dose by route and regimen on each interval; retain original updates with no return to never-used. No person-level final route is assigned backwards in time. 2. Route: oral, transdermal (plaster/gel and source mapped KUTSPRO), other formulations; R strict max cumulative source doses chooses type. Formula m51 ~ht_dosform1, m52 ~ht_dosform1+year_cat + fland_mode + udda_mode + q_ind_mode + parity_cat + dm + hypertens + afli + valv + hf + statin + many_diag + civst_mode. D10 explicitly stops route targets requiring ties/carry resolution, D04 stops future-dose/overlap-dependent assignments; do not invent tie-break variants. Summary events/time/IR and each coefficient/CI are distinct components. R551 positional first5 includes covariate rows beyond exposure; named coefficient mapping is permitted only if exact source fit/model resolved, never extra strata. 3. Regimen: estrogen-only, estrogen+cyclic progestogen, estrogen+continuous progestogen. Source estrogen+separate progestin names Jaydess/Kyleena/Mirena/Levosert become continuous, remaining separate progestin cyclic (R410); original combined G03FA/G03FB classes remain. Strict cumulative maximum as for route, D04/D10. m61 ~ht_regimen1, m62 ~ht_regimen1+A_R. Do not add dose, molecule, route-regimen cross-classification or pairwise head-to-head models. 4. Initiation age: R616 constructs first ht4-start-based object but R618 overwrites it. Actual retained recipe sets first_ht=tstart for ht==1, min across woman's exposed intervals, zero on unexposed history, cut[0,1),[.1,52),[52,57),[57,150). P labels45-51/52-56/>=57 because eligible entry is45. After checking exact first exposure and source time rounding, groups correspond to source categories. m101

~first_ht, m102 ~first_ht+A_R. The overwritten R616 is not a second authorized sensitivity; do not execute both. If prior by-reference mutation or unresolved source dose/history makes age ambiguous, stop this component. D08 adjusted attribution remains separate. 5. Export t2_route, t2_regimen, t2_initiation rows in source order with never comparator and three exposure rows per target, summary/crude/adjusted; keep route/regimen missing even if initiation rows are computed. Source exposure reference HR1 is not assigned to missing nonreference estimates. No new interaction P-values, dose-response trend test or subgroup ranking statistic.

SCIENTIFIC MEANING. Compare each stratum to its source never-use group; do not infer that differences in separate HRs establish superiority between active treatments. Unequal availability, precision, channeling and residual confounding constrain the treatment narrative. Preserve mixed patterns and observational claim ceiling, including when a single route is supported and others blocked.

Complete registration fields

```
id: atom-000004
topic: Treatment heterogeneity
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds:
- target: atom-000001
  bond: prereg
  description: Planned interpretation relationship to population mortality;
    not an input dependency or outcome-dependent activation.
preregistration:
  version: 1
  mode: canonical
  question: What associations are reported for the source treatment forms,
    progestogen regimens and initiation-age groups?
  rationale: These related exposure contrasts address heterogeneity in the
    MHT association. They stay together because they partition the same
    ↔ treatment
    history for one source question, with distinct coefficients and boundaries
    rather than a composite effect. This is independent of surgery-specific
    population selection.
  assumption: The source question is observational. Correct record linkage,
    source-equivalent histories, treatment/censoring specification and
    ↔ reported-model
    attribution are prerequisites; neither schema nor source agreement proves
    absence of confounding.
  target_strength: associational
  target_shape: association
  ratings:
    importance: 3
    novelty: 1
    feasibility: 1
    data_advantage: 3
    informativeness: 0.5
  tables:
  - BEF
  - VNDS
```

- DOD
- LMDB
- LPR_ADM
- LPR_DIAG
- LPR_OPR
- LPR_SKSOPR
- LPR_A_KONTAKT
- LPR_A_DIAGNOSE
- LPR_A_PROCREGISTRERING
- MFR
- CIV
- IND
- UDDA
- UDDF

activation: 'TRUE'

verdicts:

- key: boundary
 - region: After a completed evidence-backed assessment, no requested
 - ↪ empirical
 - component can be produced because every component is
 - ↪ source-underdetermined
 - or unavailable. Operational failure is excluded from this region.
 - says: Report precise no-estimate boundaries. There is no empirical
 - ↪ mortality
 - finding or replication-agreement conclusion.
- key: partial
 - region: At least one requested empirical component is available and at least one is explicitly unresolved, unavailable, not estimable or
 - ↪ disclosure-suppressed
 - after completed assessment. No operationally unfinished component is hidden by this state.
 - says: Report the available target-specific estimates, intervals and
 - ↪ descriptions
 - and all unavailable components. Partial coverage is not statistical confirmation or a uniform evidence direction.
- key: complete
 - region: All requested empirical components within this Atom, including code-only targets explicitly assigned to it, have resolved source
 - ↪ specifications
 - and completed reportable outputs; no unavailable component is omitted.
 - says: Report the entire source-aligned result vector, including mixed estimates and intervals. Complete coverage is not evidence of agreement, significance, equivalence, causality or successful replication.
- key: otherwise
 - region: otherwise
 - says: State the precise unresolved interpretation/coverage issue and any available component evidence without forcing it into a scientific sign or replication-success category. Unrun/operational failure remains a terminal, not this state.

protocol: |2

AUTHORITY AND SCOPE. Incorporate every applicable instruction in
 ↳ prereg/data-contract.md (schema table, ordered construction,
 ↳ inference and output contract), replication/clinical-codes.yaml and
 ↳ replication/deviations.md. These are part of this protocol, not
 ↳ optional implementation advice. P/S/R are the checked source locators
 ↳ defined in source-map.md. Only explicit source rules or lossless
 ↳ engineering may be implemented. Each STOP_UNRESOLVED_SOURCE /
 ↳ STOP_UNAVAILABLE_INPUT emits a component-specific coverage row and
 ↳ absent numerical output. A shared upstream gap propagates to all and
 ↳ only its consumers. Do not treat this stage's documented gap as a
 ↳ realized Atom result: all Atoms remain prereg and pending.

COMMON POPULATION / TIME. Women born1950-1977, alive and resident at45;
 ↳ immigration before25; no prior source contraindication, MHT or
 ↳ bilateral oophorectomy. Unit is woman's (start,stop] age interval;
 ↳ endpoint is CPR death, censoring as P3, end2023-07-31 except cause-sp
 ↳ ecific2022-12-31. Exact calendar/birthday and migration conflicts
 ↳ D01/D02, source-file mapping D03/D05 and applicable covariates
 ↳ D06/D07 must be resolved before consuming estimates. No missing key,
 ↳ date or excluded-history gap may be converted to zero/no-history. The
 ↳ source ever-use process starts unexposed and never returns to
 ↳ unexposed; source cumulative-duration process and its unresolved
 ↳ episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95%
 ↳ intervals and source software defaults in data-contract.md. P's
 ↳ two-sided .05 convention may describe individual source-specified
 ↳ tests; it does not adjudicate replication. No power calculation,
 ↳ multiplicity correction, alternative estimator, extra subgroup or
 ↳ empirical threshold is added. Use results/atom-000004/estimates.csv,
 ↳ summaries.csv, coverage.csv, adjudication.md and
 ↳ source-comparison.csv with the shared exact schemas. Do not create a
 ↳ current result file from published values. At runtime compute each
 ↳ source-defined output independently once its required inputs/rules
 ↳ are resolved; an adjusted fit's failure never overwrites
 ↳ crude/summary siblings. Retain exact target/exposure/component
 ↳ labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has termi
 ↳ nal no_result and no registered local state. For an evidence-backed
 ↳ completed module, apply verdicts in the listed first-matching order.
 ↳ Boundary/partial/complete concern coverage only. The scientific
 ↳ evidence is the entire labelled estimate/CI/descriptive record. Mixed
 ↳ signs and uncertain estimates remain mixed; missing results are not
 ↳ neutral. No local key means causal validity, significance, equivalenc
 ↳ e, replication success or independent review. Shared cohort joins and
 ↳ checks are inside this protocol, not separate experimental nodes.

SOURCE. P3-5 stratified analyses, Table2 final three blocks;
 ↳ R384-427,544-566,615-630.

EXECUTABLE PSEUDOCODE.

1. REQUIRE_SOURCE(shared eligible history, dose units and episode
 ↳ assignment). Before MHT use, all categories reference never-used.
 ↳ Keep total cumulative MHT and cumulative dose by route and regimen on
 ↳ each interval; retain original updates with no return to never-used.
 ↳ No person-level final route is assigned backwards in time.

2. Route: oral, transdermal (plaster/gel and source mapped KUTSPRO),
 ↳ other formulations; R strict max cumulative source doses chooses
 ↳ type. Formula m51 ~ht_dosform1, m52 ~ht_dosform1+year_cat +
 ↳ fland_mode + udda_mode + q_ind_mode + parity_cat + dm + hypertens +
 ↳ afli + valv + hf + statin + many_diag + civst_mode. D10 explicitly
 ↳ stops route targets requiring ties/carry resolution, D04 stops
 ↳ future-dose/overlap-dependent assignments; do not invent tie-break
 ↳ variants. Summary events/time/IR and each coefficient/CI are distinct
 ↳ components. R551 positional first5 includes covariate rows beyond
 ↳ exposure; named coefficient mapping is permitted only if exact source
 ↳ fit/model resolved, never extra strata.
3. Regimen: estrogen-only, estrogen+cyclic progestogen,
 ↳ estrogen+continuous progestogen. Source estrogen+separate progestin
 ↳ names Jaydess/Kyleena/Mirena/Levosert become continuous, remaining
 ↳ separate progestin cyclic (R410); original combined G03FA/G03FB
 ↳ classes remain. Strict cumulative maximum as for route, D04/D10. m61
 ↳ ~ht_regimen1, m62 ~ht_regimen1+A_R. Do not add dose, molecule,
 ↳ route-regimen cross-classification or pairwise head-to-head models.
4. Initiation age: R616 constructs first ht4-start-based object but R618
 ↳ overwrites it. Actual retained recipe sets first_ht=tstart for ht==1,
 ↳ min across woman's exposed intervals, zero on unexposed history,
 ↳ cut[0,.1),[.1,52),[52,57),[57,150). P labels45-51/52-56/>=57 because
 ↳ eligible entry is45. After checking exact first exposure and source
 ↳ time rounding, groups correspond to source categories. m101
 ↳ ~first_ht, m102 ~first_ht+A_R. The overwritten R616 is not a second
 ↳ authorized sensitivity; do not execute both. If prior by-reference
 ↳ mutation or unresolved source dose/history makes age ambiguous, stop
 ↳ this component. D08 adjusted attribution remains separate.
5. Export t2_route, t2_regimen, t2_initiation rows in source order with
 ↳ never comparator and three exposure rows per target,
 ↳ summary/crude/adjusted; keep route/regimen missing even if initiation
 ↳ rows are computed. Source exposure reference HR1 is not assigned to
 ↳ missing nonreference estimates. No new interaction P-values,
 ↳ dose-response trend test or subgroup ranking statistic.

SCIENTIFIC MEANING. Compare each stratum to its source never-use group;
 ↳ do not infer that differences in separate HRs establish superiority
 ↳ between active treatments. Unequal availability, precision,
 ↳ channeling and residual confounding constrain the treatment
 ↳ narrative. Preserve mixed patterns and observational claim ceiling,
 ↳ including when a single route is supported and others blocked.

Registered analysis module

These related exposure contrasts address heterogeneity in the MHT association. They stay together because they partition the same treatment history for one source question, with distinct coefficients and boundaries rather than a composite effect. This is independent of surgery-specific population selection.

All empirical results and local registered state are unfilled. The numerical design ratings are native serialization compatibility fields, not scientific evidence, feasibility validation, novelty competition or a replication score. Feasibility1 reflects explicit unresolved source/input dependencies; no independent scientific acceptance is claimed.

Reference implementation

Entry point and supported/unresolved components for `atom-000004`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.

atom-000005: Cause specific mortality

AUTHORITY AND SCOPE. Incorporate every applicable instruction in `prereg/data-contract.md` (schema table, ordered construction, inference and output contract), `replication/clinical-codes.yaml` and `replication/deviations.md`. These are part of this protocol, not optional implementation advice. P/S/R are the checked source locators defined in `source-map.md`. Only explicit source rules or loss-less engineering may be implemented. Each `STOP_UNRESOLVED_SOURCE` / `STOP_UNAVAILABLE_INPUT` emits a component-specific coverage row and absent numerical output. A shared upstream gap propagates to all and only its consumers. Do not treat this stage's documented gap as a realized Atom result: all Atoms remain `prereg` and `pending`.

COMMON POPULATION / TIME. Women born1950–1977, alive and resident at45; immigration before25; no prior source contraindication, MHT or bilateral oophorectomy. Unit is woman's (start,stop] age interval; endpoint is CPR death, censoring as P3, end2023-07-31 except cause-specific2022-12-31. Exact calendar/birthday and migration conflicts D01/D02, source-file mapping D03/D05 and applicable covariates D06/D07 must be resolved before consuming estimates. No missing key, date or excluded-history gap may be converted to zero/no-history. The source ever-use process starts `unexposed` and never returns to `unexposed`; source cumulative-duration process and its unresolved episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95% intervals and source software defaults in `data-contract.md`. P's two-sided .05 convention may describe individual source-specified tests; it does not adjudicate replication. No power calculation, multiplicity correction, alternative estimator, extra subgroup or empirical threshold is added. Use `results/atom-000005/estimates.csv`, `summaries.csv`, `coverage.csv`, `adjudication.md` and `source-comparison.csv` with the shared exact schemas. Do not create a current result file from published values. At runtime compute each source-defined output independently once its required inputs/rules are resolved; an adjusted fit's failure never overwrites crude/summary siblings. Retain exact target/exposure/component labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has terminal `no_result` and no registered local state. For an evidence-backed completed module, apply verdicts in the listed first-matching order. Boundary/partial/complete concern coverage only. The scientific evidence is the entire labelled estimate/CI/descriptive record. Mixed signs and uncertain estimates remain mixed; missing results are not neutral. No local key means causal validity, significance, equivalence, replication success or independent review. Shared cohort joins and checks are inside this protocol, not separate experimental nodes.

SOURCE. P3 outcome/covariate methods, P5–7 cause-specific narrative and Table3; S1 cause lists; R738–799. No nonfatal disease endpoint or competing-risk cumulative-incidence model is added.

EXECUTABLE PSEUDOCODE. 1. REQUIRE SOURCE(shared cohort/exposure plus underlying-cause history through2022-12-31). All-cause death remains from CPR; cause register classifies the primary cause only. Verify source date fields and duplicate/person identity before joining. R741 dar fields pnr=k_cpr,tilgrund=c_dod1,ddato parsed d_dodsdo,c_liste_49; R743 dar1 uses c_dodtilgrundl_acme, parsed d_statdato,c_listea/b. R745 rbind, order person/tilgrund, first per person. D15 stops conflicting versions/classifications and missing-cause decisions; no lexical-choice repair. 2. Cardiovascular: C_LISTEA A-08/A-09 or C_LISTE_49 23–28. Cancer: A-02 or C_LISTE_49 3,4,5,6,7,8,9,10,11,13,14, deliberately no invented12. Other: remaining source primary causes. Source code first assigns cancer then overwrites cardiovascular and missing→other. Need reconciled category/version rule before endpoint assignment. Missing or inconsistent cause is not automatically other in the translated faithful target. No substitute classification from diagnosis history. 3. End follow-up on2022-12-31 per P. Source interval-start filters with year_start<decimal_date(2023-01-01) need D09 reconciliation; no inferred exact horizon from comments or filenames. For each cause, deaths from other classified causes end follow-up with event0, own cause event1; this preserves the source cause-specific Cox estimand. R749/763 use event_pop1; R777 other uses dar2 directly. D15 gates endpoint-specific consumers independently, with no all-cause subtraction substitute. 4. Collapse source MHT duration to0/<5/>=5, keeping D04. For each cause compute events, source person time and rate per10000. m201 cardiovascular, m211 cancer, m221 other: Surv(tstart,tstop,event_cause)~ht_cat2. No Fine-Gray estimator, competing-risk risk ratio, cumulative incidence, survival curve or added cause test. Retain these source-supported crude/summary siblings if their own mapping/time prerequisites become resolved. 5. Cardiovascular adjusted R756 m202: ~ht_cat2+year_cat+fland_mode+civst_mode+udda_mode+q_ind_mode+parity_cat+dm+many_diag. R omits hyperten/afli/valv/hf/statin. T3 footnote also omits valvular, while P3 text explicitly omits only four listed mediators, and civil remains D08. No adjusted cardiovascular estimate until exact reported set reconciled. Cancer m212: ~ht_cat2+year_cat + fland_mode + udda_mode + q_ind_mode + parity_cat + dm + hypertens + afli + valv + hf + statin + many_diag + civst_mode, subject to D08. Other m222 has same adjusted formula but R785 adj22 extracts m92. D14 absolutely blocks the other adjusted target; neither m92 copying nor m222 substitution is authorized. A source-supported repair must arrive as a documented amendment before execution. 6. Source cause proportions R797–799 divide each cause count by sum of three cause counts; report cardiovascular/cancer/other death percentages only after common horizon and exhaustive mutually exclusive source categories are established. Do not force proportions to sum100 by residual filling or rounding changes. Numerator/denominator gaps remain absent in summaries.csv. 7. Export t3_cardiovascular, t3_cancer, t3_other with all3 duration rows and summary/crude/adjusted components; cause_proportions summaries. Preserve each

unavailable cell alongside any faithful sibling estimates. Published other-cause values remain reference-only under replication/, never copied to estimates.

SCIENTIFIC MEANING. Cause-specific associations describe different mortality hazards with other-cause censoring; they are not estimates of absolute competing-risk mortality probability. Component results may point in different directions or be unavailable. An all-cause estimate cannot fill a cause-specific gap, and available cancer/cardiovascular results do not resolve adjusted other mortality. No aggregate replication pass/fail or invented equivalence statement.

Complete registration fields

```
id: atom-000005
topic: Cause specific mortality
atom_result: prereg
atom_strength: null
atom_shape: null
evidence_paths: []
concepts: []
literature: []
bonds:
- target: atom-000001
  bond: prereg
  description: Planned interpretation relationship to population mortality;
    not an input dependency or outcome-dependent activation.
preregistration:
  version: 1
  mode: canonical
  question: How is the source MHT association distributed across
    ↪ cardiovascular,
    cancer and other mortality?
  rationale: Cardiovascular, cancer and other mortality jointly elaborate
    the all-cause association while sharing the cause-register horizon and
    classification. Their distinct cause-specific estimates and missing
    ↪ components
    are preserved within one original analysis module; they are not one Atom
    per cause or coefficient.
  assumption: The source question is observational. Correct record linkage,
    source-equivalent histories, treatment/censoring specification and
    ↪ reported-model
    attribution are prerequisites; neither schema nor source agreement proves
    absence of confounding.
  target_strength: associational
  target_shape: association
  ratings:
    importance: 3
    novelty: 1
    feasibility: 1
    data_advantage: 3
    informativeness: 0.5
  tables:
  - BEF
  - VNDS
  - DOD
  - LMDB
  - LPR_ADM
  - LPR_DIAG
  - LPR_OPR
  - LPR_SKSOPR
```

- LPR_A_KONTAKT
- LPR_A_DIAGNOSE
- LPR_A_PROCREGISTRERING
- MFR
- CIV
- IND
- UDDA
- UDDF
- DODSAARS
- DODSAASG
- DODSAARSAGER

activation: 'TRUE'

verdicts:

- key: boundary
 - region: After a completed evidence-backed assessment, no requested
 - ↪ empirical
 - component can be produced because every component is
 - ↪ source-underdetermined
 - or unavailable. Operational failure is excluded from this region.
 - says: Report precise no-estimate boundaries. There is no empirical
 - ↪ mortality
 - finding or replication-agreement conclusion.
- key: partial
 - region: At least one requested empirical component is available and at least one is explicitly unresolved, unavailable, not estimable or
 - ↪ disclosure-suppressed
 - after completed assessment. No operationally unfinished component is hidden by this state.
 - says: Report the available target-specific estimates, intervals and
 - ↪ descriptions
 - and all unavailable components. Partial coverage is not statistical confirmation or a uniform evidence direction.
- key: complete
 - region: All requested empirical components within this Atom, including code-only targets explicitly assigned to it, have resolved source
 - ↪ specifications
 - and completed reportable outputs; no unavailable component is omitted.
 - says: Report the entire source-aligned result vector, including mixed estimates and intervals. Complete coverage is not evidence of agreement, significance, equivalence, causality or successful replication.
- key: otherwise
 - region: otherwise
 - says: State the precise unresolved interpretation/coverage issue and any available component evidence without forcing it into a scientific sign or replication-success category. Unrun/operational failure remains a terminal, not this state.

protocol: |2

AUTHORITY AND SCOPE. Incorporate every applicable instruction in

- ↪ prereg/data-contract.md (schema table, ordered construction,
- ↪ inference and output contract), replication/clinical-codes.yaml and
- ↪ replication/deviations.md. These are part of this protocol, not
- ↪ optional implementation advice. P/S/R are the checked source locators
- ↪ defined in source-map.md. Only explicit source rules or lossless
- ↪ engineering may be implemented. Each STOP_UNRESOLVED_SOURCE /
- ↪ STOP_UNAVAILABLE_INPUT emits a component-specific coverage row and
- ↪ absent numerical output. A shared upstream gap propagates to all and
- ↪ only its consumers. Do not treat this stage's documented gap as a
- ↪ realized Atom result: all Atoms remain prereg and pending.

COMMON POPULATION / TIME. Women born1950-1977, alive and resident at45;
 ↪ immigration before25; no prior source contraindication, MHT or
 ↪ bilateral oophorectomy. Unit is woman's (start,stop] age interval;
 ↪ endpoint is CPR death, censoring as P3, end2023-07-31 except cause-sp
 ↪ ecific2022-12-31. Exact calendar/birthday and migration conflicts
 ↪ D01/D02, source-file mapping D03/D05 and applicable covariates
 ↪ D06/D07 must be resolved before consuming estimates. No missing key,
 ↪ date or excluded-history gap may be converted to zero/no-history. The
 ↪ source ever-use process starts unexposed and never returns to
 ↪ unexposed; source cumulative-duration process and its unresolved
 ↪ episode timing are fully specified in shared construction.

INFERENCE / OUTPUT. Preserve the source Cox formulas, age scale, 95%
 ↪ intervals and source software defaults in data-contract.md. P's
 ↪ two-sided .05 convention may describe individual source-specified
 ↪ tests; it does not adjudicate replication. No power calculation,
 ↪ multiplicity correction, alternative estimator, extra subgroup or
 ↪ empirical threshold is added. Use results/atom-000005/estimates.csv,
 ↪ summaries.csv, coverage.csv, adjudication.md and
 ↪ source-comparison.csv with the shared exact schemas. Do not create a
 ↪ current result file from published values. At runtime compute each
 ↪ source-defined output independently once its required inputs/rules
 ↪ are resolved; an adjusted fit's failure never overwrites
 ↪ crude/summary siblings. Retain exact target/exposure/component
 ↪ labels, model IDs and intervals; no overall positive/negative score.

ORDERED ADJUDICATION. First distinguish operational failure: it has termi
 ↪ nal no_result and no registered local state. For an evidence-backed
 ↪ completed module, apply verdicts in the listed first-matching order.
 ↪ Boundary/partial/complete concern coverage only. The scientific
 ↪ evidence is the entire labelled estimate/CI/descriptive record. Mixed
 ↪ signs and uncertain estimates remain mixed; missing results are not
 ↪ neutral. No local key means causal validity, significance, equivalenc
 ↪ e, replication success or independent review. Shared cohort joins and
 ↪ checks are inside this protocol, not separate experimental nodes.

SOURCE. P3 outcome/covariate methods, P5-7 cause-specific narrative and
 ↪ Table3; S1 cause lists; R738-799. No nonfatal disease endpoint or
 ↪ competing-risk cumulative-incidence model is added.

EXECUTABLE PSEUDOCODE.

1. REQUIRE_SOURCE(shared cohort/exposure plus underlying-cause history
 ↪ through2022-12-31). All-cause death remains from CPR; cause register
 ↪ classifies the primary cause only. Verify source date fields and
 ↪ duplicate/person identity before joining. R741 dar fields
 ↪ pnr=k_cpr,tilgrund=c_dod1,ddato parsed d_dodsdto,c_liste_49; R743
 ↪ dar1 uses c_dodtilgrundl_acme, parsed d_statdato,c_listea/b. R745
 ↪ rbind, order person/tilgrund, first per person. D15 stops conflicting
 ↪ versions/classifications and missing-cause decisions; no
 ↪ lexical-choice repair.
2. Cardiovascular: C_LISTEA A-08/A-09 or C_LISTE_49 23-28. Cancer: A-02
 ↪ or C_LISTE_49 3,4,5,6,7,8,9,10,11,13,14, deliberately no invented12.
 ↪ Other: remaining source primary causes. Source code first assigns
 ↪ cancer then overwrites cardiovascular and missing+other. Need
 ↪ reconciled category/version rule before endpoint assignment. Missing
 ↪ or inconsistent cause is not automatically other in the translated
 ↪ faithful target. No substitute classification from diagnosis history.

3. End follow-up on 2022-12-31 per P. Source interval-start filters with
 - year_start < decimal_date(2023-01-01) need D09 reconciliation; no
 - inferred exact horizon from comments or filenames. For each cause,
 - deaths from other classified causes end follow-up with event0, own
 - cause event1; this preserves the source cause-specific Cox estimand.
 - R749/763 use event_pop1; R777 other uses dar2 directly. D15 gates
 - endpoint-specific consumers independently, with no all-cause
 - subtraction substitute.
4. Collapse source MHT duration to 0/5/5, keeping D04. For each cause
 - compute events, source person time and rate per 10000. m201
 - cardiovascular, m211 cancer, m221 other:
 - Surv(tstart, tstop, event_cause) ~ ht_cat2. No Fine-Gray estimator,
 - competing-risk risk ratio, cumulative incidence, survival curve or
 - added cause test. Retain these source-supported crude/summary
 - siblings if their own mapping/time prerequisites become resolved.
5. Cardiovascular adjusted R756 m202: ~ht_cat2+year_cat+flland_mode+civst_
 - mode+udda_mode+q_ind_mode+parity_cat+dm+many_diag. R omits
 - hyperten/afli/valv/hf/statin. T3 footnote also omits valvular, while
 - P3 text explicitly omits only four listed mediators, and civil
 - remains D08. No adjusted cardiovascular estimate until exact reported
 - set reconciled. Cancer m212: ~ht_cat2+year_cat + flland_mode +
 - udda_mode + q_ind_mode + parity_cat + dm + hypertens + afli + valv +
 - hf + statin + many_diag + civst_mode, subject to D08. Other m222 has
 - same adjusted formula but R785 adj22 extracts m92. D14 absolutely
 - blocks the other adjusted target; neither m92 copying nor m222
 - substitution is authorized. A source-supported repair must arrive as
 - a documented amendment before execution.
6. Source cause proportions R797-799 divide each cause count by sum of
 - three cause counts; report cardiovascular/cancer/other death
 - percentages only after common horizon and exhaustive mutually
 - exclusive source categories are established. Do not force proportions
 - to sum100 by residual filling or rounding changes.
 - Numerator/denominator gaps remain absent in summaries.csv.
7. Export t3_cardiovascular, t3_cancer, t3_other with all3 duration rows
 - and summary/crude/adjusted components; `cause\_proportions` summaries.
 - Preserve each unavailable cell alongside any faithful sibling
 - estimates. Published other-cause values remain reference-only under
 - replication/, never copied to estimates.

SCIENTIFIC MEANING. Cause-specific associations describe different

- mortality hazards with other-cause censoring; they are not estimates
- of absolute competing-risk mortality probability. Component results
- may point in different directions or be unavailable. An all-cause
- estimate cannot fill a cause-specific gap, and available cancer/cardi
- ovascular results do not resolve adjusted other mortality. No
- aggregate replication pass/fail or invented equivalence statement.

Registered analysis module

Cardiovascular, cancer and other mortality jointly elaborate the all-cause association while sharing the cause-register horizon and classification. Their distinct cause-specific estimates and missing components are preserved within one original analysis module; they are not one Atom per cause or coefficient.

All empirical results and local registered state are unfilled. The numerical design ratings are native serialization compatibility fields, not scientific evidence, feasibility validation, novelty competition or a replication score. Feasibility1 reflects

explicit unresolved source/input dependencies; no independent scientific acceptance is claimed.

Reference implementation

Entry point and supported/unresolved components for `atom-000005`: [execution index](#). See [engineering review](#). This pointer does not amend the registered protocol.