Public Traceability of Medical Device Recalls Across US Food and Drug Administration Data Systems

Public medical device records are more useful for oversight when recalled products can be connected across data systems. We estimated this recoverability in 18,876 US Food and Drug Administration recall-product records initiated during 2019–2025, using October 2026 snapshots. The initial search identified 1,036 candidate records; source review confirmed links in 1,028. A stratified probability audit reviewed 1,233 initially unlinked records in 400 recall events. A record qualified when its narrative supported an identifier exactly matching both the Global Unique Device Identification Database and Manufacturer and User Facility Device Experience files. The audit identified 227 qualifying records, corresponding to an estimated 3,145 additional records (95% confidence interval, 2,360–3,931). Estimated full-cohort recoverability was 4,173 records (95% confidence interval, 3,388–4,959), or 22.1% (17.9%–26.3%). Intervals quantify sampling uncertainty, not classification error. Some linked records postdated recall initiation, limiting interpretation to retrospective recovery. A validated public recall–identifier crosswalk could make existing information easier to inspect and reuse. Source-based audits, correction histories, and measured maintenance costs could establish accountability for its quality. Shared identifiers do not establish recalled-lot involvement or device-related harm; operational benefits require prospective evaluation.

Keywords: *medical device recalls; unique device identification; public data; record linkage; regulatory accountability*

ONLY FOR REVIEW

Introduction

Public access to medical device safety records can support oversight only if users can identify which records concern the same product. A recall notice describes a corrective action, a device-identification record describes the product, and an adverse-event report describes a reported problem. Connecting those records can help researchers inspect the information underlying a safety concern and help health systems locate relevant product information. Yet publishing each record separately does not necessarily make those connections straightforward. The policy question is therefore not simply whether information is public, but whether its identity, context, and subsequent corrections are sufficiently clear to permit reliable reuse.

The US Food and Drug Administration (FDA) established the unique device identification (UDI) system through a 2013 final rule, creating a standardized basis for identifying devices through distribution and use (US Food and Drug Administration 2013). Policy analyses have emphasized that identifiers must also be captured and used across clinical and regulatory systems to achieve their intended value (Kinard and McGiffert 2021, Dhruva et al. 2023). This distinction matters for public accountability: an identifier can exist on a label or in a database without appearing in a form that allows an external user to connect the relevant public records.

We define public traceability narrowly as the ability to connect a recall-product record to device-identification and adverse-event files through a source-supported

identifier. A device identifier (DI) identifies a version or model, rather than an individual device unit. A shared DI therefore does not establish that a reported event involved a recalled lot or that the device caused harm. Our outcome concerns the recoverability of a documented connection between public records, not the completeness or effectiveness of medical device safety surveillance.

A failed automated search can reflect missing information, a value embedded in an unfamiliar narrative format, or an extraction rule that does not recognize otherwise usable information. Counting only records returned by a search combines these distinct possibilities. We addressed this problem by reviewing all initial search candidates and drawing a probability sample from the remaining records. The primary question was what proportion of recall-product records had a source-supported identifier matching both the Global Unique Device Identification Database (GUDID) and Manufacturer and User Facility Device Experience (MAUDE) files. We also examined explicit recall-number references and the dates of linked records to clarify what these connections could support.

This Introduction defines the policy problem and research question. The Literature Review situates the study within work on device identification and recall oversight. Materials and Methods describes the sources, review process, probability sampling, and estimation. Results presents the findings, followed by a Discussion of their implications for public-data stewardship and their limitations. Conclusions identifies the principal implication for future evaluation. The Appendix provides detailed provenance, alternative definitions, and supporting tables.

Literature Review

The UDI framework distinguishes device identity from production information. The DI is the fixed component identifying a device's labeler and specific version or model; production information can identify attributes such as lot, serial number, or expiration date. GUDID contains device-identification information, and AccessGUDID provides public access to those records (National Library of Medicine n.d.). This architecture offers a common reference for product identity, but a model-level connection must not be interpreted as proof that a particular unit or lot was involved in an event. Package identifiers also require careful treatment because a package and a contained device may have different identifiers.

Kinard and McGiffert (2021) argued that device labeling alone was insufficient and emphasized consistent identifier use in clinical records, adverse-event reports, recall alerts, and other systems. Dhruva et al. (2023) similarly assessed progress and remaining implementation challenges a decade after the UDI rule. These analyses frame identifier adoption as a coordination problem across organizations and information systems. They also motivate a narrower empirical question: how much of the information already released to the public can be connected without obtaining additional confidential data or making uncertain matches based on product names?

Mooghali et al. (2023) examined UDI availability in Class I medical device recall notices from 2018 through 2022 and found complete UDIs in 54 of 211 unique

ONLY FOR REVIEW

recalls. That study established an important distinction between the existence of an identification standard and its presence in recall communications. Our study examines a different property: whether an identifier supported by a recall narrative can be recovered in two other public data sources. The different periods, units, recall populations, and outcome definitions prevent a direct comparison of the numerical estimates. In particular, finding a complete UDI in a notice is not the same as finding a DI in both GUDID and an adverse-event file.

The structure of the underlying public files also affects interpretation. OpenFDA provides programmatic access to FDA data, but access does not remove the need to understand the originating records and their relationships (Kass-Hout et al. 2016). Ensign and Cohen (2017) described the separate components and linkage structure of MAUDE files. A device row is not interchangeable with a report, and multiple devices or record connections do not necessarily represent multiple patients or distinct harms. FDA cautions that MAUDE reports cannot establish incidence or causation and have reporting limitations (US Food and Drug Administration n.d.a). A traceability analysis must therefore preserve its unit of observation and avoid turning counts of database connections into clinical outcome estimates.

Recent government oversight reports place these technical questions in a broader policy setting. The US Government Accountability Office (2024) identified limited UDI use in electronic health records and billing claims, together with funding considerations, as challenges to active postmarket surveillance. Its subsequent review of medical device recalls described staffing constraints and difficulties coordinating recall strategies, recommending workforce planning and assessment of whether additional authority was needed (US Government Accountability Office 2025). These reports concern operational oversight and institutional capacity. They do not establish the source of any missing identifier in the public files examined here, but they show why a proposed information improvement must specify who would maintain it and what resources that work would require.

FDA's correction-and-removal reporting instructions already address the submission of identifiers, including grouping identifiers by model when several models are involved (US Food and Drug Administration n.d.d). The remaining empirical issue is whether the public representation permits reliable cross-system reuse. We investigate that representation rather than compliance by individual manufacturers or the adequacy of FDA's internal records. By separating confirmed initial links from estimated links missed by the search, the study provides a basis for evaluating a public recall–identifier crosswalk: a structured list of relationships with supporting evidence, defined identifier roles, and a correction history. Its operational benefit would require separate prospective testing.

ONLY FOR REVIEW

Materials and Methods

Design, Sources, and Units of Analysis

This cross-sectional study included all 18,876 recall-product records in the frozen openFDA device-recall export with initiation dates from January 1, 2019, through December 31, 2025. The export was updated October 2, 2026. We used the October 1, 2026, AccessGUDID full release; MAUDE annual device files for 2019–2025; and the consolidated MAUDE report file through 2025 (US Food and Drug Administration n.d.b, n.d.c). Analyses of the frozen data were completed October 4, 2026. Source versions were retained so that changes in live public databases would not alter the analytic cohort during manuscript revision.

The unit of analysis was the recall-product record, identified by cfres_id. Multiple product records could belong to the same recall event, identified by res_event_number. We distinguished product records, recall events, record–identifier pairs, and MAUDE reports throughout the analysis. A recall-product record could contain several identifiers and link to several reports, but it counted only once for each record-level outcome. This prevented products with many identifiers or reports from dominating estimates of the proportion of recall records with a recoverable link.

The study used public regulatory records and did not link patient-level data. The investigators completed Emory University's institutional determination form, which indicated that the project was neither human subjects research nor a clinical investigation and did not require institutional review board review. Informed consent was not applicable. Reporting addresses applicable items from the Strengthening the Reporting of Observational Studies in Epidemiology statement and the REporting of studies Conducted using Observational Routinely-collected health Data statement (von Elm et al. 2007, Benchimol et al. 2015). The Appendix maps relevant reporting items to their locations.

Initial Search and Candidate Review

The initial search extracted values following specified DI, UDI-DI, or Global Trade Item Number (GTIN)/DI labels in the recall code_info field. It required the same normalized identifier to appear in GUDID and the MAUDE UDI-DI field. Normalization removed whitespace and converted letters to uppercase while preserving leading zeros. Matching used identifier values rather than similarity between product or manufacturer names. GS1 check-digit validation supported interpretation of numeric identifiers (GS1 2026). The search defined a candidate queue before the probability audit; it was not designed or evaluated as an optimized retrieval system.

Claude (Anthropic), a large language model, initially screened 1,345 candidate record–identifier pairs by applying written decision rules to supplied source text. L.G. conducted the second review, and S.P. checked every case against source text, including concordant classifications, and resolved disagreements. Review distinguished the identity and role of an identifier, its supporting source context, and

ONLY FOR REVIEW

product-code agreement. Exact database membership alone did not establish that a value had been correctly interpreted in the recall narrative.

The candidate exercise included a 40-pair pilot, all of which remained in the analysis. Review criteria separated identifier confirmation from product-code agreement and retained package-identifier roles. Eight ambiguous pairs remained unconfirmed. Final criteria and the analysis protocol were documented on October 4, 2026, after review; they were not prospectively registered. Candidate concordance describes agreement within this model-assisted review process. Independence of the candidate second review from model classifications was not separately established, so that concordance is not presented as reliability between two unaided human readers.

Probability Audit of Initially Unlinked Records

The audit population comprised the 17,840 records outside the 1,036-record candidate queue, distributed across 6,010 recall events. Events were grouped by initiation year and the strongest identifier-marker family in their initially unlinked records: explicit UDI/GTIN-DI, generic DI, or no DI marker in code_info. These categories and seven years defined 21 strata. Events, rather than individual product records, were sampled to preserve the grouping of products belonging to the same recall and to account for that grouping in variance estimation.

We selected 400 events without replacement using seed 20261004 and reviewed every initially unlinked product record in the selected events, yielding 1,233 records. All 49 eligible events with generic DI markers were included. Each remaining group received at least 10 sampled events; additional events were allocated in proportion to the square root of the eligible event count. Selection probabilities therefore varied across groups. We used event-specific inverse selection probabilities to estimate totals rather than treating the raw percentage of positive sampled records as representative of the entire audit population. Table A1 provides the sampling design.

The audit packet supplied complete product-description, recall-reason, root-cause, code-information, recall-action, and distribution fields. L.G. manually read all 1,233 source narratives without seeing those classifications. S.P. then checked every concordant record, disagreement, and flagged case against source text and reconciled decisions with rule-guided model assistance. S.P.'s verification was unmasked. No separate review by an independent reference-standard reviewer was completed. These distinctions are important because masking one review does not make the entire adjudication process independent.

A record met the strict definition when its narrative supported a value as a DI, UDI, or GTIN, including a clearly labeled table or GS1 application identifier (01), and the identical normalized value appeared in both GUDID and MAUDE. Unclear source context was excluded. A visible identifier without both database memberships did not qualify. A separate 86-record diagnostic exercise contributed no probability weights and was not pooled into the estimate; 22 records that overlapped the probability sample underwent fresh source review. The Appendix

ONLY FOR REVIEW

preserves the review provenance and the distinction between identifier presence and cross-system linkage.

Outcomes and Estimation

The primary outcome was the estimated number and proportion of all 18,876 recall-product records with a recoverable cross-system identifier under the strict definition. The total combined the 1,028 confirmed candidate records with a Horvitz–Thompson estimate of additional records in the audit population (Horvitz and Thompson 1952). Full-cohort recoverability was selected as the primary presentation during manuscript revision; the original analysis emphasized recovery within the initially unlinked population. The frozen classifications were unchanged by that presentation decision. A nested secondary outcome required a GUDID Primary DI rather than allowing either primary or nonprimary identifiers.

For each sampled event, we counted the number of initially unlinked product records meeting the outcome. Within each stratum, the sum of these event totals was multiplied by the ratio of eligible to sampled events. Summing those weighted contributions yielded the estimated additional-record total. Variance was calculated from the stratum-specific variation in event totals, with finite-population correction. Sampled events with no positive records remained in the calculations, and fully enumerated strata contributed no sampling variance. Normal-approximation 95% confidence intervals were calculated before rounding. No hypothesis tests were performed.

Confirmed candidate counts were treated as fixed because all candidate pairs were reviewed. The eight ambiguous candidate records contributed zero to the strict outcome; counting all eight would increase full-cohort recoverability by 0.04 percentage points. Confidence intervals quantify sampling uncertainty conditional on recorded decisions. They do not incorporate errors in source data, interpretation, adjudication, or identifier recovery. A separate Python standard-library implementation reproduced the weighted totals and standard errors without using the original estimator; this checked the calculations, not the validity of source-context judgments.

Secondary measures included the recoverable share of initially unlinked records and the sensitivity of the evaluated initial search. Search sensitivity was the confirmed candidate count divided by the estimated full-cohort recoverable total, with interval limits obtained by transforming the audit estimate's limits. Sensitivity analyses added four UPN-labeled records under a post hoc alternative rule or all six exact-match records with unclear context. These were separate alternative definitions, not cumulative adjustments. Another outcome required an identifier-linked MAUDE report to contain the product's exact recall number. A separate full-cohort search examined recall-number references in MAUDE's removal/correction-number field.

Revision analyses described each stratum's contribution, the source formats of strict-positive records, and the timing of linked records. GUDID identifiers were mapped to parent records and their publication dates; the earliest parent date was used when several parents existed. MAUDE device-file report keys were joined to

ONLY FOR REVIEW

report receipt dates, counting each recall-record–report pair once. Publication and receipt dates were compared with recall initiation. These unweighted descriptions do not establish when a particular identifier was added or what a recall narrative contained historically. They therefore cannot establish contemporaneous public traceability.

Results

Cohort and Confirmed Candidate Records

The cohort contained 18,876 recall-product records across 6,456 recall events. The initial search identified 1,345 candidate record–identifier pairs involving 1,036 product records and 568 events. Review confirmed 1,337 pairs (99.4%), including 1,311 with a GUDID Primary DI and 26 with a nonprimary identifier. Eight pairs remained ambiguous despite exact membership in both source files. At least one confirmed pair occurred in 1,028 product records, or 5.4% of the cohort, and 566 events. Among these, 1,010 product records had a confirmed Primary DI link. Table 1 describes cohort and sample counts by initiation year; Figure 1 shows how the review components contribute to the full-cohort estimate.

Table 1. *Recall Cohort and Review Sample by Initiation Year*

Initiation year	Cohort records	Candidate records	Confirmed records	Sampled events	Sampled records
2019	3,011	3	3	55	125
2020	2,602	2	2	56	122
2021	2,260	16	9	61	209
2022	2,276	128	127	60	347
2023	2,612	283	283	54	126
2024	3,277	325	325	54	134
2025	2,838	279	279	60	170
Total	**18,876**	**1,036**	**1,028**	**400**	**1,233**

Source: Authors' analysis of frozen FDA and AccessGUDID files. Counts are recall-product records unless otherwise specified. Candidates had an extracted identifier in both GUDID and MAUDE before source review; confirmed records had at least one source-supported pair. The audit sampled recall events and reviewed every initially unlinked record in selected events. Sample counts are unweighted and annual selection probabilities differ. Initiation year may differ from the year in the recall number.

Candidate classifications were concordant for 1,342 of 1,345 pairs (99.8%). Product-code agreement occurred in 1,215 pairs (90.3%) and was assessed separately from identifier confirmation. These are process measures rather than independently established measures of coding accuracy. The initial candidate share also varied across years, but the study did not test whether changes in that share represented improvement in contemporaneous identifier publication. The data were

ONLY FOR REVIEW

collected retrospectively, and annual search yield can reflect differences in narrative format, source coverage, and later record updates.

Probability Audit and Full-Cohort Recoverability

In the probability sample, model and manual classifications agreed on identifier presence for 1,106 of 1,233 records (89.7%). After reconciliation and verification, 701 records contained a source-supported identifier, 518 did not, and 14 remained unclear. Exact matching to both databases identified 227 records meeting the strict linkage definition; 226 had a Primary DI match. Thus, the number containing an interpretable identifier was substantially larger than the number satisfying the complete cross-system outcome. Raw classification counts and corrections are reported in the Appendix.

The weighted audit estimated 3,145 additional recoverable records (95% CI, 2,360–3,931), equivalent to 17.6% of the initially unlinked population (95% CI, 13.2%–22.0%). Combining these with confirmed candidates yielded an estimated 4,173 recoverable records across the full cohort (95% CI, 3,388–4,959), or 22.1% (95% CI, 17.9%–26.3%). The Primary DI-only estimate was 4,143 records (95% CI, 3,357–4,930), or 22.0% (95% CI, 17.8%–26.1%) (Table 2). The estimated sensitivity of the initial search was 24.6% (95% CI, 20.7%–30.3%) under the strict outcome definition.

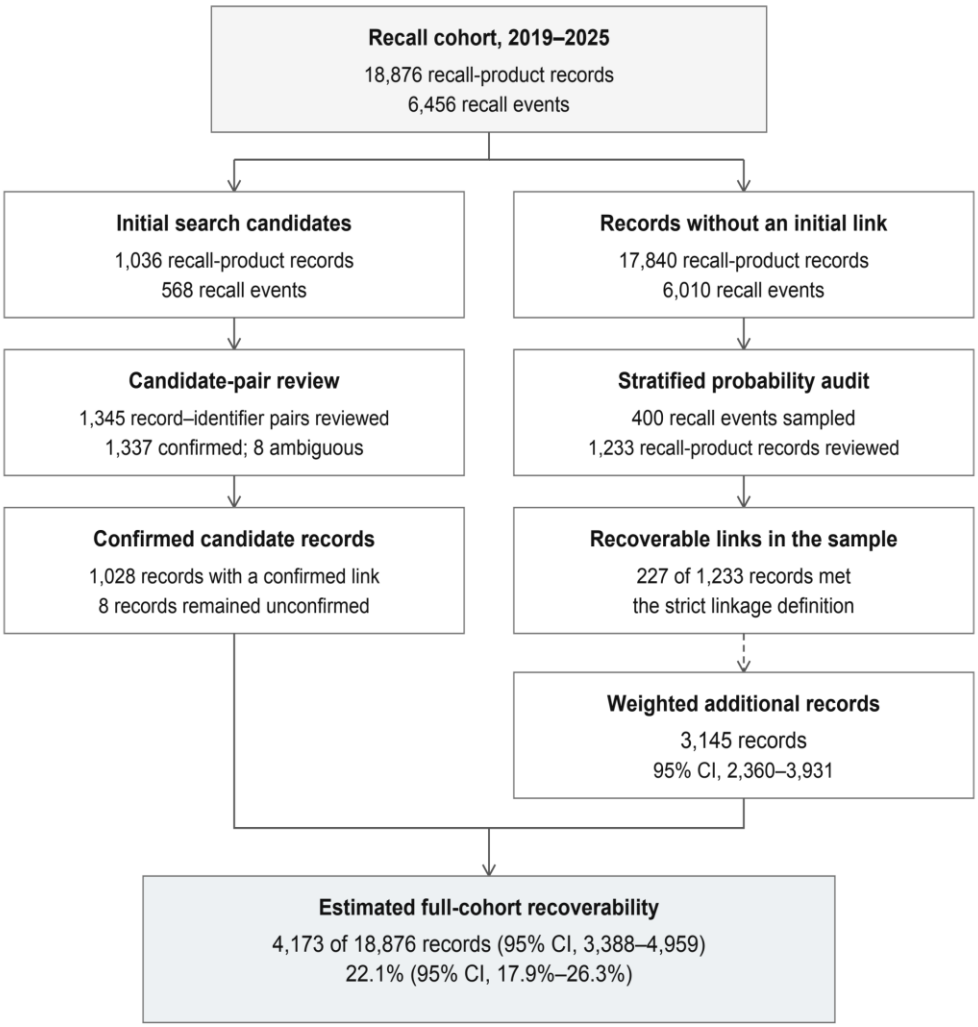
Table 2. *Estimated Full-Cohort Recoverability of Cross-System Identifiers*

Link definition	Confirmed candidates	Positive audit records	Estimated cohort records (95% CI)	Cohort share, % (95% CI)
Any GUDID identifier	1,028	227	4,173 (3,388–4,959)	22.1 (17.9–26.3)
GUDID Primary DI	1,010	226	4,143 (3,357–4,930)	22.0 (17.8–26.1)

Source: Authors' design-based estimates. Denominator: 18,876 recall-product records. Total = confirmed candidates + weighted additional records in the 17,840-record audit frame. Each record counts once. Any-identifier linkage includes primary and nonprimary GUDID identifiers; Primary DI is a nested outcome, so rows must not be added. Eight ambiguous candidate records remain unconfirmed. Confidence intervals reflect sampling uncertainty, not classification error; calculations precede rounding. DI: device identifier; CI: confidence interval.

ONLY FOR REVIEW

Figure 1. *Cohort Construction and Estimated Full-Cohort Recoverability*



Source: Authors' analysis. Solid arrows show observed review steps and combination of components; the dashed arrow shows estimation using inverse event-selection probabilities. Recall events can appear in both branches; branch event counts are not additive. Eight candidate records remained unconfirmed. The lower box combines 1,028 confirmed records with the estimated 3,145 additional records. Counts and intervals are rounded after calculation. CIs quantify sampling uncertainty conditional on the classifications.

Strata without a DI marker in code_info contributed an estimated 744 additional records, representing 23.7% of the audit estimate and 32.3% of its sampling variance. Adding four UPN-labeled records increased the audit estimate to 3,198; including all six unclear-context exact-match records yielded 3,215. A post hoc inventory of the 227 strict-positive sample records identified 57 using a labeled table or paired list, 167 using a direct label or identifier-only list, and three using GS1 (01) notation without a prose label. These formats describe the frozen positive decisions and do not provide independent validation of them.

Timing and Explicit Recall References

GUDID parent-record publication postdated recall initiation for 23 of 1,337 confirmed candidate record–identifier pairs (1.7%) and 10 of 961 strict-positive audit pairs (1.0%). MAUDE receipt dates postdated recall initiation for 45.5% of candidate record–report pairs and 52.4% of audit pairs. Median receipt-minus-initiation differences were −55 days (interquartile range, −526 to 238) and 35 days (−406 to 528), respectively. A report could occur in several recall-record pairs, so these distributions do not count unique adverse events. Table A6 provides the denominators and additional timing measures.

Among the 227 strict-positive sampled records, 15 had an exact recall-number reference in the same identifier-linked MAUDE report. This corresponded to an estimated 174 records in the audit population (95% CI, 65–283), or 1.0% (95% CI, 0.4%–1.6%). Separately, 810 of all 18,876 recall-product records (4.3%) had an exact recall-number reference somewhere in the eligible report file. These outcomes use different frames and should not be added. Neither measures the completeness of recall references among all clinically related reports.

A subsequent broader automated extraction yielded exact GUDID–MAUDE links in 1,878 cohort records (9.9%), including 83 with a same-report recall-number reference. That set was not fully adjudicated. Its higher retrieval count therefore cannot replace the 1,028 confirmed candidate records in the initial-search sensitivity calculation. The result demonstrates dependence on extraction rules and supplies a direction for future method development, but it is not a second validated estimate of full-cohort recoverability.

Discussion

Principal Findings and Interpretation

Approximately one in five recall-product records had a publicly recoverable identifier matching both GUDID and MAUDE under the study's strict definition. The estimated 22.1% comprised the 5.4% of cohort records confirmed through the initial search and an estimated additional 16.7 percentage points recovered through the probability audit. The practical finding is that a substantial portion of the recoverable connections was not found by the evaluated search. A search result alone would therefore have understated the information recoverable through source review within the specified public files.

The full-cohort estimate and initial-search sensitivity answer different questions. The former describes the amount of recoverable cross-system information under this study's rules and snapshots. The latter describes how well one extraction method found that information. Neither is a measure of FDA surveillance sensitivity. Likewise, records not meeting the strict outcome cannot all be classified as publication failures: they may lack an applicable identifier, lack a corresponding report, contain information outside the supplied fields, or require

ONLY FOR REVIEW

another valid linkage method. This distinction is necessary for interpreting both the estimate and any accountability measure derived from it.

The audit adds information beyond a larger convenience review because it includes initially unlinked records with known selection probabilities. Records from strata without a DI marker accounted for almost one-quarter of estimated additional recovery. Restricting review to narratives already recognized as containing a marker would have excluded this contribution. However, probability sampling addresses selection into the audit, not misclassification within it. The resulting uncertainty intervals should be read alongside the absence of an independent reference standard and the documented dependence of final decisions on the investigators' rules.

Our findings complement work on identifier availability in recall notices and implementation across systems (Mooghali et al. 2023, Dhruva et al. 2023). They do not imply that differing published percentages describe a temporal trend. The unit here is a recall-product record, and the outcome requires membership in two additional databases. More broadly, the results separate three tasks that are sometimes combined: identifying a product, documenting that a report references a recall, and establishing whether a particular recalled device contributed to harm. Only the first two were examined, and an identifier match by itself establishes neither the third nor the timing of public availability.

A Public Crosswalk as a Policy Response for Evaluation

The absence of a dedicated DI, UDI, or GTIN field in the analyzed recall export creates a practical publication problem when identifiers remain embedded in narrative text. This finding concerns that export, including its nested fields, rather than FDA's internal records or other endpoints. FDA's reporting instructions already describe ways to submit identifiers and group them by model (US Food and Drug Administration n.d.d). A policy response could therefore begin by evaluating how existing submitted information is represented publicly, while separately investigating information that is absent or inapplicable. The present study cannot locate an omission at the submission, processing, or public-release stage.

A structured recall–identifier crosswalk could publish each supported relationship as a separately inspectable entry. At minimum, an entry could include the recall-product record, identifier value, identifier role, source field or document, validation status, and date of verification. Primary, package, and contained-device roles should remain distinguishable. The original source text should be accessible alongside the structured interpretation so that users can determine why the relationship was accepted. Such a resource would make reuse possible without requiring each investigator or health system to repeat the same narrative interpretation.

The crosswalk should also separate device identity from recall-specific evidence. An exact identifier match could be labeled as a shared product identity, while a separate field could indicate whether the linked report contains an explicit recall-number reference. Lot or serial information, when present, would require its own validation. Presenting these relationships in separate fields would limit the risk that users interpret a model-level connection as an adjudicated harm event. It would

ONLY FOR REVIEW

also allow a corrected interpretation to change one relationship without silently rewriting the meaning of every connection involving the product.

Responsibilities, Corrections, and Measurable Accountability

Accountability requires an assigned process for correcting records as well as a published dataset. Manufacturers and other submitters could verify the model and package associated with each submitted identifier. FDA could check format, database resolution, and consistency with the submitted product description, then return unresolved discrepancies for correction. Health systems could evaluate whether the published values can be matched to purchasing or implantation records. Researchers and other users could report disputed links through a documented channel. These are proposed responsibilities for evaluation; the study did not measure current performance of these activities.

A correction history would make changes observable. Each disputed entry could retain the original value, the reason for review, the decision, and the date on which a corrected version became available. An unresolved status should remain visible rather than being represented as a confirmed link. Where identifiers are absent, an applicability field could distinguish information not required or not relevant from information expected but missing. Responsibility for resolving each category should be specified before reporting quality measures, because the same blank field can arise from different circumstances and require a different response.

Quality assessment should evaluate both published and initially unlinked records. A probability audit of published links could estimate the proportion supported by source evidence and identify errors in identifier roles. A separate probability audit of unlinked records could estimate additional recoverable information and indicate which formats the publication process misses. Both assessments should use rules established before review, preserve sampling probabilities, and report uncertainty. Independent source review would be particularly valuable for checking concordant automated and investigator decisions, a limitation that remains in the current study.

Further measures could include time from a reported discrepancy to a documented resolution, the number and type of corrections, and staff time per resolved case. Reporting these measures together would discourage treating link volume as the sole sign of success. A system that publishes more links while introducing unsupported relationships would not necessarily improve the usability of safety information. Conversely, temporarily marking a disputed link as unresolved could improve accuracy while lowering apparent coverage. An evaluation should therefore make accuracy, coverage, timeliness, and maintenance burden visible as separate outcomes.

Implementation costs cannot be inferred from the retrospective number of recoverable records. Some connections may be straightforward to structure, whereas long tables, multiple package levels, or conflicting source fields may require manual investigation. The workforce concerns identified in government oversight reports make this distinction consequential (US Government Accountability Office 2024, 2025). A prospective pilot could measure the work

ONLY FOR REVIEW

required to validate and maintain a defined set of new recall records before expanding to older data. Costs and achievable correction times should inform the design rather than being assumed negligible because the original files are publicly available.

Temporal Availability and Future Evaluation

The timing analyses limit what this study implies for recall response. Some GUDID records were published after recall initiation, and many linked MAUDE reports were received afterward. Even a publication or receipt date preceding initiation does not establish that the relevant identifier was present at that time; later supplements may have added information. A retrospective crosswalk could support investigation and historical review while still being unsuitable as evidence of what a hospital, manufacturer, or regulator could identify when a recall began. Prospective evaluation would require archived versions and explicit dates for individual field changes.

FDA announced the Adverse Event Monitoring System in March 2026 and described plans to incorporate legacy adverse-event data; its device-file documentation also describes the planned availability of MAUDE data in that system (US Food and Drug Administration 2026, n.d.c). Such changes provide an opportunity to evaluate identifier publication and version histories prospectively. The legacy-file results presented here do not assess the new system. A useful evaluation would test whether validated identifiers are available at recall issuance and whether they reduce the time or effort needed to identify affected products, while separately assessing false matches and correction costs.

For that evaluation, the relevant denominator should be specified in advance. One measure could ask whether a recall record contains a validated product identifier when such an identifier is applicable. Another could ask whether that identifier connects to an adverse-event report already public on the recall date. These measures should remain separate because an accurate product identifier need not have a corresponding adverse-event report. A third measure could assess whether the published lot or serial information permits identification of the affected units. The current study supplies a retrospective estimate for its own cross-system definition; it does not supply baseline values for all three prospective measures.

Repeated snapshots would also help distinguish new information from improved extraction. If an identifier appears in a later crosswalk, an audit should record whether the underlying source changed, a previously overlooked value was recovered, or an earlier interpretation was corrected. Otherwise, increasing coverage could be mistaken for improving contemporaneous reporting when it reflects retrospective enrichment. Preserving those distinctions would make future evaluations more informative and would allow external users to reproduce what was available at a specified time. The feasibility and workload of maintaining that history would themselves be outcomes of the proposed pilot.

ONLY FOR REVIEW

Limitations

First, recoverability is conditional on the selected public sources, fields, identifier rules, and snapshots. It does not encompass every valid product relationship. Exact matching may miss differently represented values, and even correctly matched values may be mislabeled in the source. Requiring a corresponding MAUDE identifier also means that records without such a report cannot satisfy this outcome, irrespective of the quality of the recall notice. The estimated complement of recoverability should not be described as a rate of defective reporting, missing device identities, or missed harms.

Second, the review process limits certainty about classification. Claude generated initial classifications, and S.P., who helped develop the rules, conducted the final verification. L.G.'s manual probability-sample review was masked to model classifications, but reconciliation was unmasked and no independent reference-standard review was completed. Corrections included concordant responses, illustrating why high agreement does not establish accuracy. The exact Claude version and authenticated execution times were unavailable. Criteria were documented after review, pilot records were retained, and the full-cohort emphasis was selected during revision. These features are disclosed rather than treated as prospective validation.

Third, confidence intervals include sampling uncertainty but not classification uncertainty or source-data error. The 22 records overlapping the earlier diagnostic review and the eight unconfirmed candidate records introduce additional interpretive qualifications. Alternative context rules examine a small set of specified ambiguous cases; they do not bound all possible errors. Reproduction of the weighted calculations establishes numerical consistency conditional on the decisions and does not replace an independent review of the underlying narratives.

Finally, phased UDI implementation and exceptions complicate interpretation of absent identifiers (US Food and Drug Administration 2013, n.d.e). Model and package matches do not establish patient exposure, a recalled lot, causal harm, or when an organization learned about a problem. MAUDE lacks a denominator of device use, and repeated record connections must not be interpreted as incidence. No policy intervention, reporting compliance, institutional workload, or patient outcome was evaluated. The findings support testing public-data publication and quality assurance; they do not establish regulatory inaction, preventable harm, or the effectiveness of the proposed crosswalk.

Conclusions

An estimated 22.1% of medical device recall-product records contained a publicly recoverable identifier matching both device-identification and adverse-event files in the analyzed snapshots. Most of that estimated recovery was identified through the probability audit rather than the initial search. A validated public recall–identifier crosswalk could make existing connections easier to inspect and reuse. Its value should be evaluated through source-based accuracy checks, audits of unlinked

UNDER PEER REVIEW

records, visible correction histories, and measured maintenance costs. These are concrete responsibilities for public-data stewardship. Prospective evaluation is needed to establish whether the connections are available when recalls begin and whether they improve operational decisions.

References

Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, RECORD Working Committee (2015) The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLOS Medicine 12(10): e1001885. https://doi.org/10.1371/journal.pmed.1001885

Dhruva SS, Ross JS, Wilson NA (2023) Unique device identifiers for medical devices at 10 years. JAMA Internal Medicine 183(10): 1045–1046. https://doi.org/10.1001/jamainternmed.2023.3572

Ensign LG, Cohen KB (2017) A primer to the structure, content and linkage of the FDA's Manufacturer and User Facility Device Experience (MAUDE) files. eGEMs 5(1): 12. https://doi.org/10.5334/egems.221

GS1 (2026) GS1 general specifications standard. Release 26.0, January 2026. Available at: https://ref.gs1.org/standards/genspecs/26.0.0/. Accessed 30 October 2026.

Horvitz DG, Thompson DJ (1952) A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47(260): 663–685. https://doi.org/10.1080/01621459.1952.10483446

Kass-Hout TA, Xu Z, Mohebbi M, Nelsen H, Baker A, Levine J, Johanson E, Bright RA (2016) OpenFDA: an innovative platform providing access to a wealth of FDA's publicly available data. Journal of the American Medical Informatics Association 23(3): 596–600. https://doi.org/10.1093/jamia/ocv153

Kinard M, McGiffert L (2021) Medical device tracking—how it is and how it should be. JAMA Internal Medicine 181(3): 305–306. https://doi.org/10.1001/jamainternmed.2020.7797

Mooghali M, Ross JS, Kadakia KT, Dhruva SS (2023) Availability of unique device identifiers for Class I medical device recalls from 2018 to 2022. JAMA Internal Medicine 183(7): 735–737. https://doi.org/10.1001/jamainternmed.2023.0727

National Library of Medicine (n.d.) About AccessGUDID. Available at: https://accessgudid.nlm.nih.gov/about-gudid. Accessed 30 October 2026.

US Food and Drug Administration (2013) Unique device identification system: final rule. Federal Register 78(185): 58786–58828. Available at: https://www.govinfo.gov/content/pkg/FR-2013-09-24/pdf/2013-23059.pdf.

US Food and Drug Administration (2026) FDA launches new adverse event look-up tool. News release, 11 March. Available at: https://www.fda.gov/news-events/press-announcements/fda-launches-new-adverse-event-look-tool. Accessed 30 October 2026.

US Food and Drug Administration (n.d.a) About Manufacturer and User Facility Device Experience (MAUDE) database. Available at: https://www.fda.gov/medical-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-facilities/about-manufacturer-and-user-facility-device-experience-maude-database. Accessed 30 October 2026.

US Food and Drug Administration (n.d.b) Device recall overview. openFDA. Available at: https://open.fda.gov/apis/device/recall/. Accessed 30 October 2026.

ONLY FOR REVIEW

US Food and Drug Administration (n.d.c) MDR data files. Available at: https://www.fda.gov/medical-devices/medical-device-reporting-mdr-how-report-medical-device-problems/mdr-data-files. Accessed 30 October 2026.

US Food and Drug Administration (n.d.d) Medical device recalls and reports of corrections and removals, Part 7, Part 806: frequently asked questions. Available at: https://www.fda.gov/media/170726/download. Accessed 30 October 2026.

US Food and Drug Administration (n.d.e) Unique device identification system (UDI system). Available at: https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/unique-device-identification-system-udi-system. Accessed 30 October 2026.

US Government Accountability Office (2024) Medical devices: FDA has begun building an active postmarket surveillance system. GAO-24-106699. Washington, DC: US Government Accountability Office. Available at: https://www.gao.gov/products/gao-24-106699.

US Government Accountability Office (2025) Medical device recalls: HHS and FDA should address limitations in oversight of recall process. GAO-26-107619. Washington, DC: US Government Accountability Office. Available at: https://www.gao.gov/products/gao-26-107619.

von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLOS Medicine 4(10): e296. https://doi.org/10.1371/journal.pmed.0040296

ONLY FOR REVIEW

Appendix. *Data Provenance and Additional Analyses*

This Appendix accompanies the manuscript and is included in the complete submission file. GUDID: Global Unique Device Identification Database; MAUDE: Manufacturer and User Facility Device Experience; DI: device identifier; UDI: unique device identification; GTIN: Global Trade Item Number; CI: confidence interval.

A1. *Additional Source and Parsing Details*

The source files were the openFDA device-recall export updated October 2, 2026; AccessGUDID full release dated October 1, 2026; DEVICE2019.txt through DEVICE2025.txt; and mdrfoiThru2025.txt. An openFDA device-enforcement export updated September 28, 2026, supported exploratory recall-hazard classification; the enforcement and recall endpoints were not pooled. No National Practitioner Data Bank or Open Payments records were linked.

The initial MAUDE source audit enumerated 15,733,007 annual-device rows. Its product-code analysis excluded 412 malformed-width rows and 1 row with no product code, leaving 15,732,594. This product-code audit is distinct from identifier-membership verification: the latter read a row if its UDI-DI column was available. Of 15,704,402 distinct device-file report keys in the source audit, 15,704,368 matched the consolidated report file; 34 did not. The consolidated file had 6,681 physical lines with a nonstandard field count. These diagnostics describe file parsing; they are not counts of patients or harms.

The October 4 data-lock manifest preserves source/output hashes. The evidence archive retains a hash-checked copy of inputs used for numerical re-derivation. The frozen recall cohort comprises 18,876 cfres_id records and 6,456 res_event_number events. An event can contain both candidate and initially unlinked product records; the branch event counts are not additive.

A2. *Review Provenance and Ambiguous Decisions*

An exploratory search generated 1,069 pairs across 801 records. Expanded labels added 276 pairs, yielding the evaluated queue of 1,345 pairs across 1,036 records. Claude generated initial source-context classifications stored under S.P.'s investigator identifier; these entries were not unaided human classifications. Candidate second-review masking was not established. All 40 pilot pairs remained; three overall pilot classifications differed over product-code discrepancies. Criteria separated identifier confirmation from product-code agreement and retained package roles. A rule accepted unambiguous typographic UDI-DI variants when checksum-valid and resolved in GUDID, while retaining the original uncertain response.

For the probability sample, L.G. manually read all 1,233 source narratives without model classifications. S.P. verified every record against source text and reconciled decisions unmasked with model assistance. Raw model/L.G. decisions were 586 Yes/Yes, 520 No/No, 108 Yes/No, 15 Unclear/No, and 4 Unclear/Yes.

ONLY FOR REVIEW

Raw concordance was 1,106/1,233 (89.7%). Records 197505 and 197506 were initially No in both passes but were corrected after source verification. Final presence classifications were 701 Yes, 518 No, and 14 Unclear; strict cross-system linkage occurred in 227 records.

Six of the 14 Unclear records had exact membership in both files. Four were UPN-labeled records (186284, 194639, 195965, 195966); two used a nonstandard GUIDID label or an unlabeled table. Strict linkage excluded all six. Post hoc Rule 4 added the four UPN cases, and expanded context added all six. These are alternative definitions, not additive adjustments. No independent reference-standard validation was completed.

The 86-record diagnostic exercise had 82 concordant and four reconciled presence decisions but no probability weights. Second-pass independence was not established. Twenty-two records overlapped the probability audit and received instructions for fresh review without prior decisions. All 22 had identifier presence in both exercises. Two met strict linkage in the audit: 206606 and 210355. For 210355, review recovered a contained-device DI in addition to the package DI considered earlier. Presence status and strict linkage are distinct variables.

A3. *Estimation and Reproducibility*

For each stratum, multiply the sum of qualifying product-record counts in sampled events by the ratio of eligible to sampled events. Sum these contributions to obtain the estimated additional-record total, T. The stratum variance contribution equals the squared eligible-event count multiplied by one minus the sampling fraction, multiplied by the sample variance of event totals, and divided by the number of sampled events. Sum contributions over noncensus strata. Zero-outcome events are included. Census strata contribute zero sampling variance.

For strict any-identifier linkage, T = 3,145.4977 and SE(T) = 400.7814. Adding 1,028 confirmed candidates and dividing by 18,876 gives 22.1101% (95% CI, 17.9485%–26.2716%). Rounding occurs after calculation; the lower limit is 17.9%. Initial-search sensitivity is C/(C + T), with reversed transformed interval limits for T. Primary DI calculations use C = 1,010 and the corresponding audit estimate. Independent implementation checks verify arithmetic conditional on the frozen decisions; they do not verify coding accuracy.

A4. *Additional Retrieval and Revision Analyses*

The recall-reference search normalized product recall numbers and matched them to the MAUDE removal/correction-number field. The same-report outcome additionally required the recall reference in the exact report linked by identifier. The full-cohort reference count and audit-frame same-report estimate cannot be added. A subsequent broader extraction found 1,878 records with exact membership in GUDID and MAUDE, including 83 same-report recall references; this set was not fully adjudicated.

A separate 397-record source-coding exercise ended with 313 Yes and 84 No decisions after 78 differences were reconciled by S.P. Its second-pass downstream

ONLY FOR REVIEW

lookups coincided with the automated reference, and a hidden machine-reference tab was present in the review materials. S.P. confirmed checking every case and disagreement. These results document process provenance, not independent validation or a census of broader extraction results.

Exploratory post hoc Rule 4 estimates among class-specific initially unlinked records were 3.8% (95% CI, 1.2%–6.5%) for regulatory device Class I, 20.1% (14.8%–25.4%) for Class II, and 38.0% (2.2%–73.8%) for Class III. Class III included 43 sampled records; these estimates do not establish a reliable class difference. Regulatory device class differs from recall hazard classification. Expanded-context record–report connections were highly concentrated: ten recall records accounted for 83.5% of 127,200 connections. These duplicated connections are not unique adverse events or patients.

For GUDID timing, the earliest parent publication date was used for identifiers with multiple parents (one candidate pair and 14 audit pairs). All pairs had valid dates. For MAUDE timing, all 415,837 selected report keys matched metadata, with no malformed selected metadata rows. Across device files, 387 physical lines without a readable UDI-DI column were skipped. Receipt/publication dates do not establish when an identifier was added or what the recall narrative contained at initiation.

The Codex-assisted post hoc source-format inventory described the 227 frozen strict-positive audit records without recoding outcomes. Labeled tables or paired lists can also contain GS1 notation. The three records classified as GS1-only were 203895, 203897, and 203898. A typographic UI-DI header in record 213817 remained a labeled paired list under the original accepted decision. The inventory is not independent human validation.

ONLY FOR REVIEW

Table A1. *Stratified Event Sampling Design*

Year	Marker family	Eligible events	Sampled events	Eligible records	Sampled records	Event weight
2019	Explicit	209	21	533	33	9.952
2019	Generic	3	3	8	8	1.000
2019	None	757	31	2,467	84	24.419
2020	Explicit	186	20	500	41	9.300
2020	Generic	5	5	21	21	1.000
2020	None	723	31	2,079	60	23.323
2021	Explicit	278	23	626	77	12.087
2021	Generic	9	9	35	35	1.000
2021	None	610	29	1,583	97	21.034
2022	Explicit	544	28	1,482	115	19.429
2022	Generic	11	11	22	22	1.000
2022	None	217	21	644	210	10.333
2023	Explicit	621	29	1,728	72	21.414
2023	Generic	5	5	11	11	1.000
2023	None	169	20	590	43	8.450
2024	Explicit	693	30	2,383	59	23.100
2024	Generic	5	5	10	10	1.000
2024	None	147	19	559	65	7.737
2025	Explicit	673	30	2,161	101	22.433
2025	Generic	11	11	29	29	1.000
2025	None	134	19	369	40	7.053
Total		**6,010**	**400**	**17,840**	**1,233**	—

Source: Authors' sampling frame. Strata combine initiation year and strongest marker family. Event weight = eligible/sample events; every initially unlinked product record in each sampled event was reviewed. Generic-marker groups were fully enumerated. Seed: 20261004. Weights shown are rounded; calculations use exact ratios.

ONLY FOR PEVIEW

Table A2. *Primary and Alternative Outcomes in Initially Unlinked Records*

Definition	Positive records	Estimated additional records (95% CI)	Audit-frame share, % (95% CI)	Search sensitivity, % (95% CI)
Strict any identifier	227	3,145 (2,360–3,931)	17.6 (13.2–22.0)	24.6 (20.7–30.3)
Strict Primary DI	226	3,133 (2,347–3,920)	17.6 (13.2–22.0)	24.4 (20.5–30.1)
Rule 4 any identifier	231	3,198 (2,410–3,985)	17.9 (13.5–22.3)	24.3 (20.5–29.9)
Rule 4 Primary DI	230	3,185 (2,397–3,974)	17.9 (13.4–22.3)	24.1 (20.3–29.6)
Expanded any identifier	233	3,215 (2,427–4,003)	18.0 (13.6–22.4)	24.2 (20.4–29.8)
Expanded Primary DI	232	3,203 (2,414–3,991)	18.0 (13.5–22.4)	24.0 (20.2–29.5)

Source: Authors' design-based estimates. Denominator: 17,840 initially unlinked records. Rule 4 adds four UPN-labeled cases; expanded context adds six unclear-context exact-match cases, including those four. The definitions are separate scenarios. Search sensitivity uses each definition's corresponding confirmed candidate count (1,028 any identifier; 1,010 Primary DI).

Table A2 (continued). *Same-Report Recall-Number Outcome*

Definition	Positive records	Estimated records (95% CI)	Frame share, % (95% CI)
Strict	15	174 (65–283)	1.0 (0.4–1.6)
Rule 4	15	174 (65–283)	1.0 (0.4–1.6)
Expanded context	16	181 (72–291)	1.0 (0.4–1.6)

Source: Authors' analysis. A qualifying record has an identifier-linked report containing its exact product recall number. Denominator: 17,840 initially unlinked records. This is not the fraction of all clinically related events with a recall reference.

ONLY FOR REVIEW

Table A3. *Review and Verification Provenance*

Component	Unit and size	Review or verification	Result and interpretation
Candidate queue	1,345 pairs; 1,036 records	Claude screening; L.G. second review; S.P. source verification of every case, including concordant cases	1,337 confirmed pairs; 8 ambiguous pairs
Candidate agreement	1,345 pairs	Model-assisted coding-pass concordance; not reliability between 2 unaided human readers	1,342 concordant classifications (99.8%)
Nonmatch diagnostic sample	86 records	Model-generated initial coding; S.P. verification; second pass	82 agreements; 4 adjudicated differences; no probability estimate
Weighted candidate-negative audit	1,233 records; 400 events	Claude screening; L.G. manual reading masked to model labels; S.P. unmasked source verification of all records and reconciliation	1,106 concordant raw decisions (586 Yes/Yes; 520 No/No); final 701 Yes, 518 No, 14 Unclear
Weighted calculation check	Frozen record-level results	Separate Python standard-library estimator	Strict, expanded, Rule 4, and same-report estimates reproduced
Later source-coding exercise	397 records	Model-generated initial coding; second-pass reference dependence not excluded; S.P. verification	313 Yes; 84 No after 78 adjudicated differences; not a census
Broader automated extraction	18,876-record cohort	Automated extraction and exact membership	1,878 GUDID–MAUDE linked records; 83 same-report full-chain records; not fully adjudicated

Source: Authors' analytic archive and review provenance. S.P.: Snehal Patel; L.G.: Logan Glosser. Confirmed masking applies to L.G.'s probability-sample review only. Model–human concordance, source verification, file membership, and numerical reproduction are different checks. No independent reference-standard validation was completed.

ONLY FOR REVIEW

Table A4. *Selected Observational and Routinely Collected Data Reporting Locations*

Reporting topic	Location	Content
Design and setting	Title, Abstract, Methods	Cross-sectional public-record linkage audit; 2019–2025 recalls; October 2026 sources
Eligibility and selection	Methods; Figure 1; Materials and Methods; Appendix A1–A2	Recall-product cohort, candidate complement, event-cluster probability sample
Data sources and linkage	Methods; Materials and Methods; Appendix A1–A2	Frozen source versions, identifiers, units, normalization, exact matching
Validation and bias	Methods; Limitations; Appendix A2; Table A3	Pilot retained; Claude screening; L.G. masked to model classifications in the probability sample; S.P. final verification; post-review criteria documentation
Study size and sampling	Methods; Table 1; Table A1	21 strata; 400 sampled events; 1,233 reviewed records
Variables and outcomes	Methods; Materials and Methods; Appendix A2–A4	Strict identifier recovery; Primary DI; same-report recall reference
Statistical methods and uncertainty	Methods; Appendix A3; Table 2	Event weights, finite-population correction, normal CIs, transformed sensitivity
Missing and ambiguous data	Methods; Results; Materials and Methods; Appendix A1–A2	8 unconfirmed candidate records; 14 Unclear audit decisions; no imputation
Sensitivity analyses	Results; Table A2	Post hoc Rule 4 and expanded context distinguished from primary outcome
Interpretation and generalizability	Discussion and Limitations	Full-cohort retrospective recovery under the strict definition; no surveillance sensitivity, harm incidence, or policy effect
Source and code access	Data Availability; Appendix A1 and A3	Public-source provenance; retained analysis package; no patient-level linkage

Source: Authors' reporting map. Applicable STROBE and RECORD domains are covered; this is a regulatory-record audit, not a patient cohort. Clinical demographic, exposure-effect, and follow-up items are not applicable.

ONLY FOR REVIEW

Table A5. *Stratum Contributions to the Weighted Audit Estimate*

Year and marker	Positive sampled events	Positive sampled records	Estimated additional records	Share of total, %	Share of variance, %
2019 Explicit	11	15	149.3	4.7	0.7
2019 Generic	0	0	0.0	0.0	0.0
2019 None	5	11	268.6	8.5	11.4
2020 Explicit	10	22	204.6	6.5	4.0
2020 Generic	3	14	14.0	0.4	0.0
2020 None	1	1	23.3	0.7	0.3
2021 Explicit	13	34	411.0	13.1	19.9
2021 Generic	3	8	8.0	0.3	0.0
2021 None	6	15	315.5	10.0	18.3
2022 Explicit	11	16	310.9	9.9	4.8
2022 Generic	3	6	6.0	0.2	0.0
2022 None	1	1	10.3	0.3	0.1
2023 Explicit	12	22	471.1	15.0	22.9
2023 Generic	1	2	2.0	0.1	0.0
2023 None	4	4	33.8	1.1	0.1
2024 Explicit	11	13	300.3	9.5	4.4
2024 Generic	3	3	3.0	0.1	0.0
2024 None	5	12	92.8	3.0	2.1
2025 Explicit	13	23	516.0	16.4	11.0
2025 Generic	4	5	5.0	0.2	0.0
2025 None	0	0	0.0	0.0	0.0
Total	**120**	**227**	**3,145.5**	**100.0**	**100.0**

Source: Authors' analysis. Strict any-identifier outcome. Add 1,028 confirmed candidates to the weighted additional total for full-cohort recovery. Zero-outcome events remain in variance calculations; census strata contribute no sampling variance. Variance shares use estimated variance 160,625.8 records². Percentages may differ from 100% because of rounding.

ONLY FOR REVIEW

Table A6. *Timing of Linked Records Relative to Recall Initiation*

Measure	Confirmed candidates	Strict-positive audit
Recall-product records, No.	1,028	227
Recall-record–identifier pairs, No.	1,337	961
Unique identifiers, No.	1,142	912
GUDID publication after initiation, pairs, No. (%)	23 (1.7)	10 (1.0)
Records with ≥1 post-initiation GUDID pair, No.	23	8
Recall-record–MAUDE-report pairs, No.	484,559	125,362
Unique MAUDE reports, No.	356,938	97,183
Report receipt before initiation, pairs, No. (%)	263,295 (54.3)	59,633 (47.6)
Report receipt on initiation, pairs, No. (%)	555 (0.1)	56 (<0.1)
Report receipt after initiation, pairs, No. (%)	220,709 (45.5)	65,673 (52.4)
Receipt minus initiation, median (IQR), d	−55 (−526 to 238)	35 (−406 to 528)
Records with earliest receipt on/before initiation, No. (%)	881 (85.7)	167 (73.6)
Records with earliest receipt after initiation, No. (%)	147 (14.3)	60 (26.4)

Source: Authors' unweighted timing analyses. A report can link to multiple recall records and appear in both groups. All selected pairs had valid dates. Earliest parent publication date was used for multiple GUDID parents. Dates do not establish when identifiers, supplements, or recall narratives became public. IQR: interquartile range.

Table A7. *Source Formats of Strict-Positive Audit Records*

Source format	Records, No. (%)	Interpretation
Labeled table or paired list	57 (25.1)	A header assigns the identifier's role among multiple fields; includes compact slash-delimited lists.
Direct label or identifier-only list	167 (73.6)	A DI, UDI, or GTIN label identifies a value or single-field list.
GS1 (01) without a prose label	3 (1.3)	Encoded GTIN notation supports the identifier role.
Total	**227 (100.0)**	**Frozen strict-positive records.**

Source: Authors' Codex-assisted descriptive inventory. Categories describe source formats, not independent validation or mutually exclusive evidence pathways; some tables also contain GS1 syntax. Frozen strict outcomes were unchanged. Ambiguous records excluded from strict linkage are not included.

UNDER PEER REVIEW