Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Representation of New HCPCS Product Codes in Public Medicare Fee-for-Service Data

Andreas Peters

a.peters81@icloud.com

Rosenheim Technical University of Applied Sciences

Research Article

Keywords: Medicare, HCPCS, administrative claims data, public-use files, health informatics, data governance, healthcare financing, measurement, health services research

Posted Date: September 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10792502/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Representation of New HCPCS Product Codes in Public Medicare Fee-for-Service Data

Andreas Peters (corresponding author)
Rosenheim Technical University of Applied Sciences, Rosenheim, Germany
Email: a.peters81@icloud.com

Research Article · *BMC Health Services Research*

Abstract

Background: Health services research on new medical technologies usually begins with items already present in released Medicare claims, conditioning the sample on an earlier administrative step. A Healthcare Common Procedure Coding System (HCPCS) Level II code makes an item nameable in a claim, yet code effectiveness alone puts no row into a public file. No prior study we identified measures how often a new product code is represented at all in released data.
Methods: Retrospective cohort of code events. All 103 additions in the official April 2021, 2022 and 2023 alpha-numeric HCPCS files entered a single-author descriptor screen; 49 supplied-product codes were retained and one S-series code removed, leaving 48. Each became effective on 1 April, giving a common service period of seven quarters through 31 December of the following year. Three annual datasets, for physician services, equipment and supplies, and outpatient care, were queried with the exact code, giving 288 code by service year by dataset cells. Representation meant at least one national exact-code row in a file covering services furnished in the effective year or the year after, assessed on 22 August 2026. Wilson intervals are descriptive; the cohort is enumerated, not sampled.
Results: Nineteen of 48 code events were represented (39.6%, Wilson reference interval 27.0 to 53.7). Eleven were present in the effective-year file, eight only in the following-year file and 29 in neither; one present in the earlier file was absent from the later. In the earlier file holding a row the physician file carried 11 codes, the equipment file eight and the outpatient file none. No exact-code row appeared in any of 144 pre-effective-year cells. Across the reference analysis and 14 alternative analyses the estimate ranged from 27.6% to 45.0%, lowest after removing the wound and skin-substitute class.
Conclusions: Code effectiveness and representation in these three datasets are separate administrative states. A missing national row is unresolved, not an absence of use, coverage or payment, and health services research that samples from released claims inherits that state, so the representation state belongs beside the coding event in the study record.

Keywords Medicare; HCPCS; administrative claims data; public-use files; health informatics; data governance; healthcare financing; measurement; health services research

1 Background

Administrative claims are the main data infrastructure of health services research on new medical technologies in the United States, and much of that work shares one starting point. A technology is chosen, its code or set of codes is identified, and use is then measured in released data. Device utilisation across specialties has been measured from procedures identifiable by distinct Current Procedural Terminology or Healthcare Common Procedure Coding System (HCPCS) codes [1]; a coding scheme has been built to separate durable medical equipment and supplies within claims [2]; remote monitoring of cardiac implants [3], remote patient monitoring in traditional Medicare [4] and seventeen years of video capsule endoscopy [5] have all been described in the same way. Every one of those designs presupposes something it cannot itself establish, namely that the code of interest reaches the released file.

Several administrative steps sit between an effective code and a released row. A service must be furnished and billed, the claim must fall within a domain that the Centers for Medicare & Medicaid Services (CMS) publishes, the record must survive aggregation and public-release rules, and the file must be released. Failure at any of those steps produces the same observable result, an absent row, whatever happened clinically. A sample built from codes already present in a released file is therefore conditioned on the outcome of the reporting process, and the codes that never reach a file are precisely the ones such a sample cannot see. Whether that conditioning matters is an empirical question about how large the unrepresented part is.

Effectiveness of a HCPCS Level II code does not settle the question either. CMS treats a code as a billing identifier and states that its existence determines neither coverage nor non-coverage [6]. Coding, coverage, payment and use are separate institutional acts with separate timing, a separation examined in work linking device regulation to reimbursement [7], in evidence that Medicare coverage decisions follow their own logic after authorisation [8], and in cross-country comparison of authorisation and payment arrangements for the same devices [9].

1.1 Adjacent research and what it leaves open

Quantitative work on new medical technology in Medicare falls into six strands, each of which takes a different administrative act as its starting event and a different endpoint as its object of measurement. Several groups reconstruct the time from Food and Drug Administration authorisation to a Medicare coverage milestone for defined technology sets

[10, 11, 12]. Other work treats coverage itself as an act with its own evidence base and its own timing, nationally and locally [8, 13, 14]. A third strand relates payment policy to diffusion, covering cochlear implants [15], coronary stents [16], positron emission tomography under a global budget [17], add-on and supplemental payment programmes [18, 19, 20] and deep brain stimulation [21]. Utilisation is then measured directly once a code exists, in video capsule endoscopy over seventeen years [5], in remote patient monitoring [4], in remote monitoring of cardiac implants [3], across specialties for procedures identifiable by distinct codes [1], and through a scheme built to separate equipment and supplies within claims [2]. Identification forms a fifth strand, in which unique device identifier projects have documented how much item identity disappears when a claim retains a procedure but not the implanted product [22, 23, 24, 25, 26, 27, 28, 29]. Comparative work makes a sixth, setting out how systems assess, authorise and pay for the same device [30, 7, 9].

Table 1 sets the six strands beside one another on the two features that matter here, the event a design begins from and the endpoint it measures. Read that way, the strands divide cleanly. Studies of authorisation, coverage and cross-system arrangement work from institutional records and never touch released claims, so representation plays no part in them. Studies of payment policy and of utilisation work from released claims and therefore begin with technologies that are already represented, which makes representation an assumption of the design instead of an object of it. Work on device identification comes closest to the present question and still asks a different one, namely whether a claim that does exist carries enough information to name the item, not whether a claim for a newly coded item reaches the released file at all.

Table 1. Where the present study sits among adjacent research strands. The two columns that matter are the event a design begins from and the endpoint it measures; the last column states whether representation of the code in released public claims is assumed by the design or measured by it.

Strand	Representative work	Starting event	Endpoint measured	Representation in released claims
Authorisation to coverage	[10, 11, 12]	FDA authorisation of a device	interval to a Medicare coverage milestone	not involved; coverage records are the source
Coverage determination	[8, 13, 14]	an authorised device under review	national or local coverage decision and its evidence base	not involved
Payment policy and diffusion	[15, 16, 17, 18, 19, 20, 21]	a payment rule, programme or reform	utilisation or spending after the rule	assumed; the technology is already represented
Utilisation of an established code	[5, 4, 3, 1, 2]	an existing, stable code	counts, trends and spending over time	assumed; the sample is drawn from released rows
Device identification in claims	[22, 23, 24, 25, 26, 27, 28, 29]	a claim that names a procedure but not the item	feasibility of carrying item identity into records and claims	addressed as specificity of an existing row, not as presence of a row
Cross-system assessment	[30, 7, 9]	a device under assessment	how systems authorise, assess and pay	not involved
Present study	—	establishment of a new product code on 1 April	presence of a national exact-code row in three specified annual files	measured directly, as the outcome

We cite representative work by strand rather than exhaustively. The table classifies designs by what they measure, not by their quality or their contribution.

The gap follows from the arrangement. No strand takes the establishment of a product code as its starting event, and none takes presence of that code in released public claims as its endpoint. To our knowledge, and on a targeted rather than systematic search of the strands above, no study quantifies whether a code that has become usable is represented in a public file at all. Its size is not a matter of curiosity, because two of the six strands sample from exactly the population that lies on the far side of it.

1.2 Research question and scope

Our question is narrow, and stating it precisely is half the contribution. Posed this way it belongs to health informatics and not to technology assessment: what is measured is a property of the released data, and the products enter only as the objects the codes name. For a cohort of product codes newly established in an April HCPCS update, what proportion is represented by at least one national exact-code row in the annual public Medicare fee-for-service files covering services furnished during the effective year or the year after it, as assessed on a single retrieval date? Four subsidiary questions follow. Representation may arise in the file for the effective service year or only in the file for the year after, so the annual state in which it arises is worth separating. Representation need not persist from one annual file to the next, so persistence is a second question rather than an assumption. Rows may sit in any of three claims domains, so which domain carries them is a third. And an estimate from 48 codes may depend on how the cohort boundary was drawn, so its stability under administrative and product reclassification is a fourth.

Three questions lie outside the design and are not asked. When CMS first released a row is not observable from annual files queried retrospectively, so no publication lag is estimated. Whether a product was used, covered, paid for or reached a beneficiary is not observable from an absent row, so no such quantity is inferred. And why any individual row is missing cannot be settled by these data, so the mechanisms

are enumerated and left unapportioned.

Direct measurement is what we report. Using the official April HCPCS updates for 2021 to 2023, we assembled a cohort of newly established supplied-product codes and answered the question above, together with its four subsidiary parts.

Three consequences make the answer relevant beyond the specific cohort, and each falls in a different part of health services research. Cohort construction in technology-use research inherits whatever the reporting process has already filtered, so the size of the unrepresented group bounds how much a claims-first sample can miss; that is a measurement problem for health informatics, and it applies to any analysis that treats a public claims aggregate as a population register. Surveillance and reproducibility turn on the same gate, since a study that cannot distinguish low use from an unreleased row cannot be replicated against a later file release without recording which state applied; the remedy is a matter of data governance and stewardship, and the closing section states it as a rule. Evaluation of reimbursement policy meets the same difficulty, because a reform that shortens one institutional interval need not shorten the others, and an evaluation that collapses coding, coverage, payment and public release into a single date can credit or blame the wrong institution; healthcare financing, healthcare governance and the design of health system reform all rest on being able to tell those events apart.

What the study examines is the data infrastructure of one payer, not clinical practice. No clinical question is asked, no treatment is compared, and nothing here touches clinical decision-making at the bedside. Its bearing on healthcare delivery and access stays indirect and precise: research on delivery and on access that samples technologies from released claims inherits the gate measured here, so the estimate constrains what such research can conclude without saying anything about delivery or access itself. Analogous under-ascertainment can arise wherever aggregated claims are published under suppression or partition rules, so the measurement problem is not confined to one payer, although the estimate reported here is specific to Medicare and to three datasets.

2 Methods

2.1 Design, unit of analysis and reporting frame

A question of this kind can be answered badly in two ways, by settling the unit of observation after seeing the data and by leaving the frame it is drawn from open. Both stay fixed here before anything is retrieved. We constructed a retrospective national cohort of administrative code events and queried it on 22 August 2026. The unit of analysis is a newly established five-character HCPCS Level II code event, not a product, manufacturer, claim, patient, provider or coverage decision. Repeated measurement cells are defined by code, service year and dataset. We claim no completed STROBE or RECORD checklist, because both instruments presuppose individual-level records that a code-event design does not have [31, 32]. Additional file 10 instead locates the source definition, the cohort construction rule, the provenance of every variable, the handling of empty and unresolved values, and the cleaning steps applied. One analysis script holds the source files, screening categories, datasets, service-period rule, exact-match rule and outcome definition. We registered no timestamped protocol, so this description reports a workflow and claims no preregistration.

2.2 Source population and cohort construction

Every addition in the official April 2021, April 2022 and April 2023 alpha-numeric HCPCS files formed the source population [33, 34, 35]. An addition had to carry action code A, an addition date of 1 April in the relevant year and a matching active effective date, which produced 21 source rows for 2021, 36 for 2022 and 46 for 2023, or 103 in total.

A second step screened long descriptors for a discrete supplied product. Eligible descriptors covered stand-alone devices, prosthetics, orthotics, diagnostic readers, software devices, components, accessories, recurring device supplies, and wound or skin-substitute products. The screen excluded drugs, biologicals, diagnostic radiopharmaceuticals, professional or facility services, model payments, procedures, tests and modifiers. We describe retained codes throughout as *supplied-product codes*, a deliberately operational label naming what the controlling CMS descriptor names and nothing else. The screen assigns no regulatory classification to any product, which matters most for the wound and skin-substitute items, where prefix alone establishes neither device status nor biological status. Screening retained 49 codes and excluded 54: 40 as drugs, biologicals or diagnostic radiopharmaceuticals, seven as services or model payments, five as procedures or tests and two as modifiers. Re-reading all 54 excluded descriptors produced one change. C9782 describes a blinded clinical-trial procedure whose intervention arm is a transcatheter intramyocardial cell transplantation, and the first pass had placed it in the drug, biological and cell-product class while placing its structurally identical sibling C9783 among the procedures; C9782 is now recorded as a procedure. The code was excluded either way, so the retained set and every result stand unchanged, and the single value that moved is logged in the corrections file of Additional file 8. One further code, S1091, was removed. Exclusion rests on one pre-defined ground, the Medicare reporting universe: codes in the S series are established for private-payer and other non-Medicare reporting rather than for Medicare use [6]. Coverage indicator I on the April record for S1091, which in the record layout marks an item as not payable by Medicare, is corroborative metadata and not the decision rule. Two retained codes, A4560 and A7049, carry coverage indicator S, a different value denoting an item non-covered by Medicare statute. Both stay in the cohort, and deliberately so. Cohort membership follows what a descriptor names, and excluding codes on a coverage classification would put a plausible determinant of the outcome into the denominator, which is the error the study exists to avoid. A sensitivity analysis restricted to coverage indicator C removes both and shows what their retention costs. After the single removal, the analytic cohort contained 48 code events (Fig. 1, panel A).

A fixed order of questions governed the screen, applied to the controlling long descriptor and to nothing else. Does the

descriptor name a modifier, a professional or facility service, a model payment, a procedure or a test? If so the code is excluded. Does it name a drug, a biological, a cell product or a diagnostic radiopharmaceutical? If so the code is excluded. Does it name a discrete supplied product, whether a stand-alone item, a component, an accessory, a recurring supply or a wound or skin-substitute product? If so the code is retained. Any code reaching the last question without a clear answer would carry a flag; none did. The single author applied that order alone. No duplicate screening was undertaken and no inter-rater statistic is available, so the classification should be read as a judgement recorded in an auditable ledger rather than as a validated instrument. Every screened row, its assigned class and the reason for its disposition are supplied in Additional file 1, which allows the boundary to be moved and the analysis repeated.

One retained code, A4100, is a not-otherwise-specified skin-substitute code rather than a code for a single named product. Retaining it follows the same operational rule as every other retention, since the descriptor names a supplied product, and its outcome enters the ledger like any other; a reader who regards a not-otherwise-specified code as a different object can identify and remove it from Additional file 2. Wound and skin-substitute items needed a separate rule. Nineteen retained codes describe such products, seven with an A prefix and twelve with a Q prefix, and a descriptor alone establishes no common regulatory status for either prefix. All 19 were therefore held together as one product class in the principal cohort and removed together in a strict-product sensitivity analysis, which avoids choosing one prefix as a device proxy after seeing the outcome.

Choice of the April cycles follows from the service-period rule. An effective date of 1 April yields the same seven-quarter service period, from 1 April of the effective year through 31 December of the following year, for every code in every cohort year; cycles with January, July or October effective dates would give service periods of eight, six and five quarters respectively, and pooling them would compare codes whose files cover unequal exposure. April 2023 is the most recent cycle for which the archived retrieval holds annual files covering both service years; it contains files for service years 2020 through 2024 and no later year. No catalogue snapshot was preserved, so the statement describes the archived query set and not the catalogue as it stood on the retrieval date.

As an independent routing check, exact code strings were sought in the final second-biannual non-drug or durable-equipment decision documents preceding each April update [36, 37, 38]. Thirty-six of the 49 screened codes were located in the corresponding document. The remaining thirteen, comprising the twelve Q-coded wound products and S1091, arose through another coding route and appear in the quarterly file without appearing in those documents. We analysed the 36-code subset separately. Location of a code string in a decision document establishes routing only; it is not a product-level regulatory adjudication.

2.3 Baseline administrative variables

Screening decides which codes enter the cohort; the entry record fixes what the April file says about each of them at the moment it becomes usable. The April files supplied the controlling long descriptor, effective date, Berenson-Eggers Type of Service category, type-of-service values, pricing indicator, coverage indicator and Outpatient Prospective Payment System fields, the last of which were empty for all 48 codes. The long descriptor supplied three product groups, namely stand-alone device or related finished item, component, accessory or recurring supply, and wound or skin-substitute item. Codes tied to one named system, or standing in an express parent-and-component relation, additionally received a product-family identifier, giving 41 families among the 48 codes. Full assignments and their basis are in Additional files 2 and 6.

Raw administrative fields were not renamed as policy outcomes. Coverage indicator C denotes carrier judgement in the record layout, not affirmative Medicare coverage; indicator I marks an item as not payable by Medicare and indicator S an item non-covered by statute. Pricing indicator 99 means that a value was not established, while 00 means that the service is not separately priced by Part B, an entry that can arise for more than one reason. CMS states in the note accompanying its annual HCPCS update that an absent coverage or payment indicator on a new code may simply mean that no national determination or fee schedule amount exists yet, and is an indicator of neither coverage nor non-coverage [39]. Neither field shows that a claim will or will not be paid. For one exploratory contrast, pricing indicators 00 and 99 were grouped as *no separately specified Part B method or value in the baseline file* and all other pricing indicators as *a specific method*.

2.4 Public data sources and query rule

Three annual CMS geography-and-service datasets, the public reporting layer of Medicare fee-for-service and a standard source for health services research, were queried through their official application programming interfaces [40, 41, 42]. Table 2 states, for each dataset, the service years used, the claims domain, the exact HCPCS field, the geography rule, the partitions that legitimately produce more than one national row, the meaning of an empty result and the limits that follow [43, 44, 45]. For a code effective in year Y, the files for service years Y and $Y + 1$ were used, which gives 288 code by service year by dataset cells across the 48 codes. Dataset identifiers, complete request URLs, retrieval date, byte count and SHA-256 hash for every archived response are listed in Additional file 7.

Table 2. Public Medicare fee-for-service datasets queried, with the rule applied to each

	Physician and Other Practitioners	Durable Medical Equipment, Devices & Supplies	Outpatient Hospitals
Service years used	2020–2024	2020–2024	2020–2024
Claims domain	physician and other practitioner services	equipment, devices and supplies furnished through suppliers	hospital outpatient services
Exact code field	`HCPCS_Cd`	`HCPCS_Cd`	`HCPCS_Cd`
Geography rule	`Rndrng_Prvdr_Geo_Lvl` equal to National	`Rfrg_Prvdr_Geo_Lvl` equal to National	`Rndrng_Prvdr_Geo_Lvl` equal to National; rule never exercised, no row returned
Partitions within one national code-year	place of service (`Place_Of_Srvc`)	supplier rental status (`Suplr_Rentl_Ind`)	none observed
Volume fields present, never pooled	beneficiary, service and rendering-provider counts	beneficiary, claim, service, supplier and referring-provider counts	not observed in this cohort
Empty result means	no national exact-code row in that file; not a clinical zero	no national exact-code row in that file; not a clinical zero	no national exact-code row in that file; not evidence of no hospital use
Known limits	annual and aggregate; Original Medicare fee-for-service only; low-volume records subject to public-release suppression	same, and eligibility depends on supplier billing rather than the site of care	same, and channel eligibility for these codes was not established

Field names are taken from the archived responses and the query rule implemented in the analysis script; dataset scope follows the CMS methodological documentation for each file. No universal suppression threshold is asserted, because none was established from the archived material.

Queries used the exact five-character HCPCS value. A returned object was accepted only when its code field matched the requested code after case and whitespace normalisation. Rows were classified from the source geography field as national or state; state-only rows were retained for quality control and did not satisfy the outcome. More than one national row for the same code, service year and dataset was permitted, since place of service and supplier rental status partition a national aggregate legitimately, and beneficiary, provider, supplier, claim and service-unit fields were never summed across partitions or datasets because their denominators and units differ. Empty responses, state-only responses and responses without a national exact-code row were assigned one common state, *no national row in the specified public files*.

Each code was also queried in the same three datasets for service year $Y - 1$, giving 144 further cells. A positive exact-code result there would have indicated code reuse, reactivation, or an error in cohort or source assignment, and would not have counted as early uptake. Separately, termination and action-effective fields in the April 2025 full HCPCS reference file were checked against each code's service-period end [46]. The accompanying CMS record layout defines the termination date as the last date on which a code may be used by Medicare providers, so the check tests whether a code remained usable throughout its assigned service period and records any successor code without transferring an outcome between predecessor and successor.

2.5 Outcome definitions

Everything to this point fixes what we asked of the data. What remains is to say exactly what a returned row must look like before it counts. For code i, service year y and dataset d, representation was defined as

$$R_{iyd} = \mathbb{1}\{\exists\, r \in \mathcal{R}_{iyd} \text{ with } \operatorname{norm}(r_{\text{code}}) = i \text{ and } r_{g_d} = \text{National}\}, \tag{1}$$

where $\mathcal{R}_{iyd}$ is the set of rows returned by the query for code i against the file for service year y in dataset d, norm strips case and whitespace, and g_d is the geography field of that dataset, `Rndrng_Prvdr_Geo_Lvl` for the physician and outpatient files and `Rfrg_Prvdr_Geo_Lvl` for the equipment file. A response that returned successfully with an empty row set gives $R_{iyd} = 0$; a query failing to complete or to parse would not enter the tally at all, and none failed. The primary outcome was

$$R_i = \max_{y \in \{Y_i, Y_i+1\}} \max_{d \in \{P,D,O\}} R_{iyd}, \tag{2}$$

where P, D and O denote the physician, equipment and outpatient datasets. Since every effective date was 1 April, the two files jointly cover services furnished across seven calendar quarters.

Wording matters here and is fixed for the remainder of the article. A file for service year y records services furnished in y; querying it on 22 August 2026 shows the state of that file on that date. Nothing in the design reveals when CMS first released a row, so no publication lag, no time to public availability and no day-level event time can be derived from these data. Annual files support exactly three states: a national row in the file for the effective service year; no row there and a row in the file for the following service year; or no row in either. Representation is also not necessarily absorbing, because a code present in one annual file may fall below a release threshold in the next. Kaplan–Meier, Cox and interval-censored survival models were therefore not fitted. Secondary outcomes were the four-state persistence pattern across the two files and the dataset combination in the first file showing a row.

2.6 Statistical analysis

With the endpoint fixed as one binary state per code, the arithmetic stays deliberately plain. We calculated counts and proportions for the analytic cohort overall and by effective year, with Wilson 95% intervals [47]. No sample-size

calculation was performed, because every addition in the three selected April files was enumerated and the cohort is therefore complete for the frame it defines rather than sampled from it. An interval of that kind is a descriptive binomial reference interval computed under a working model of independent codes. Product families violate that model, which is why a family-level estimate is reported beside it, and neither interval should be read as sampling uncertainty around a population of future coding actions. No population-weighted utilisation rate is computable for codes that have no public denominator.

Descriptive stratification and sensitivity analysis serve different purposes here. Four descriptive strata carry no adjustment and count as exploratory throughout: cohort year, product group, coverage indicator and baseline pricing group, together with the distribution of national rows across datasets and annual files. Sensitivity analyses restricted the cohort to coverage indicator C; to the 36 codes located in the corresponding second-biannual decision documents; to those decision-document codes with indicator C; to codes outside the wound and skin-substitute class; and to codes whose termination date falls after their service-period end. A family-level analysis counted a product family as represented when at least one member had a national row. Leave-one-cohort-year-out estimates test whether one April release carries the result.

One exploratory contrast compares baseline pricing categories within the 46 coverage-C codes and repeats the comparison after removal of all wound and skin-substitute items. Counts and proportions are reported for it; no test statistic appears in the main text, because a purpose-built finite cohort of administrative categories with small dependent cells offers no superpopulation for one to describe. Interval estimates for the same contrast are printed in Additional file 10 for readers who want them. Pricing indicator, code prefix, product family, type of service and Berenson-Eggers group encode overlapping administrative structure that 19 represented codes cannot separate, so no multivariable model was fitted and the words determinant, driver and effect are not used for that contrast. Statistical calculations were made in the deterministic build script, and interpretive weight follows effect size, interval width and sensitivity rather than a threshold on p [48, 49].

2.7 Verification and reproducibility

A measurement claim is worth no more than its trail. The 103-row screen, the 48-code ledger, the 288-cell outcome matrix, the 144-cell negative-control matrix, the code-lifecycle table, the product-family mapping and both source manifests accompany the manuscript as Additional files 1 to 13. Each manifest row carries the request URL, the retrieval date, the byte count and the SHA-256 hash of the stored response, so any reader who re-queries CMS can test their own download against ours; the response payloads themselves are held for deposit and are not part of the submitted package. Structural validation required 48 unique code events, six outcome cells and three negative-control cells per code, exact HCPCS matches in every non-empty response, and agreement between the code-level summary and the cell matrix. Four cells contained more than one legitimate national partition; the binary rule counted each once and added no volume fields. A4238 and E2102 received descriptor revisions during follow-up and remained linked by the unchanged exact code. Tables and figures are regenerated from the validated ledgers without manual numerical entry.

3 Results

3.1 Cohort construction and baseline structure

Of the 103 April additions, 49 supplied-product codes passed the screen and, after removal of S1091, left 48 code events: eight effective in 2021, 19 in 2022 and 21 in 2023 (Fig. 1, panel A; Table 3). Eighteen describe a stand-alone device or related finished item, eleven a component, accessory or recurring supply, and nineteen a wound or skin-substitute product. Coverage indicator C applies to 46 codes and indicator S to two. Twenty codes carry pricing indicator 00 or 99 and 28 a specific method in the baseline record. Panel B of Fig. 1 sets out the measurement design, including the pre-effective-year control.

Composition shifts across the three cycles in a way that matters for every comparison that follows. All eight 2021 codes are K-codes for stand-alone items or their supplies, and all eight carry pricing indicator 99; not one wound or skin-substitute product entered in that cycle. Wound and skin-substitute items appear from 2022 onward and account for nine of the 19 codes added in 2022 and ten of the 21 added in 2023, which is also where the specific pricing-method indicator becomes common. Cohort year and product class are therefore entangled by construction, and no analysis in a cohort of this size can separate them.

Product families group the 48 codes into 41 units. Six families hold more than one member: two 2021 neurostimulation systems each pairing a device with its monthly supply code, two 2022 systems, one 2022 compression system with three codes, and one 2023 intraurethral drainage system with two. Additional file 6 lists every family assignment and the descriptor text on which it rests. Composition of that kind sets how far any single restriction can move the estimate, which the sensitivity analyses later test.

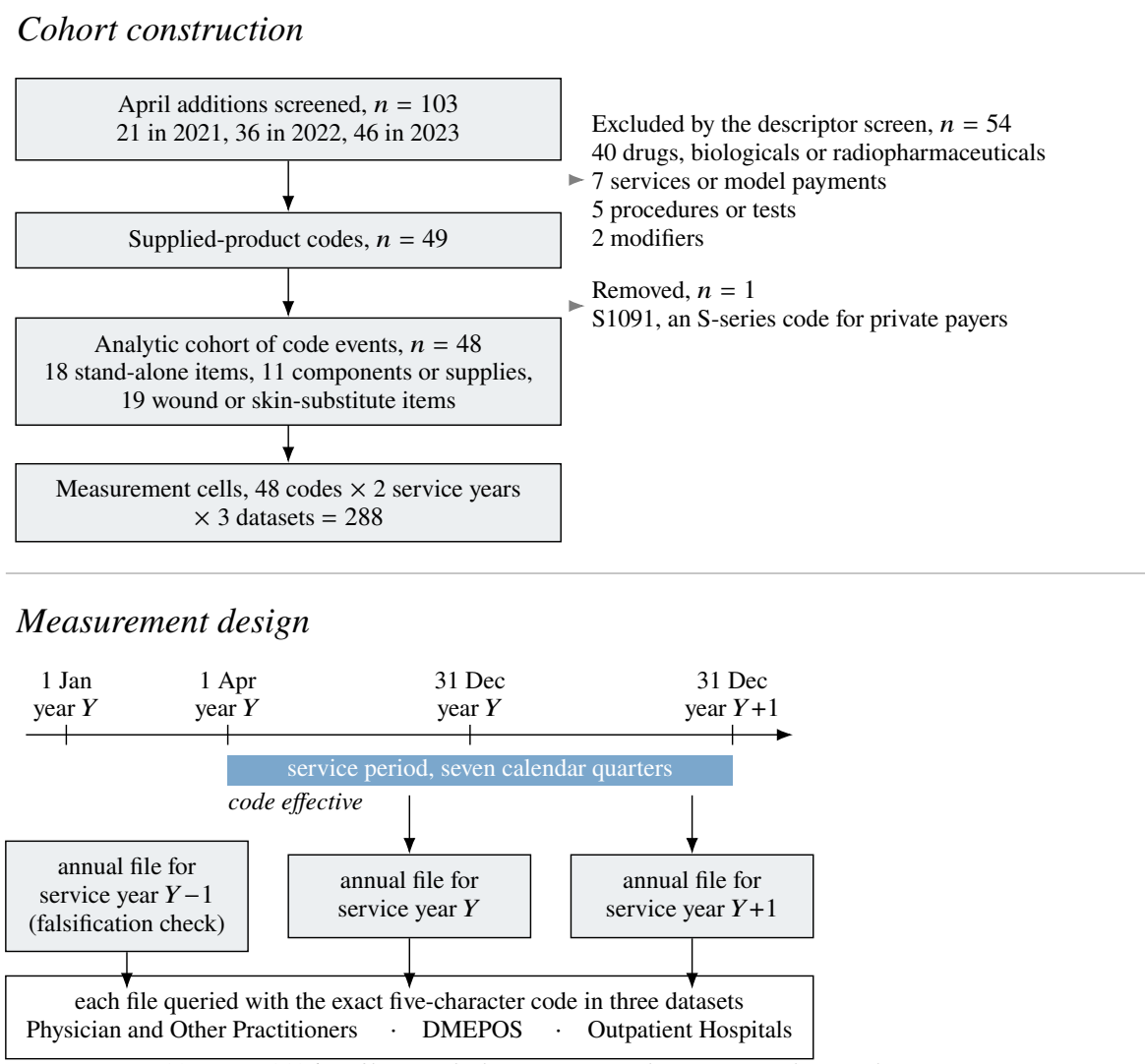


Figure 1. Cohort construction and measurement design. Panel A traces 103 April additions through the descriptor screen and the single removal to 48 code events and 288 measurement cells. Panel B shows the common service period of seven calendar quarters, the two annual files that cover it, the pre-effective-year control file and the three datasets queried for each file. All queries were run on 22 August 2026; the design records the service period a file covers and does not observe when CMS released a row.

Table 3. Cohort construction and baseline administrative characteristics

	Overall	2021	2022	2023
April additions screened	103	21	36	46
Device and other non-drug product codes	49	9	19	21
Analytic cohort after removal of S1091	48	8	19	21
Stand-alone device or related finished item	18	5	7	6
Component, accessory or recurring supply	11	3	3	5
Wound or skin-substitute item	19	0	9	10
Coverage indicator C	46	8	19	19
Coverage indicator S	2	0	0	2
Pricing indicator 00 or 99	20	8	8	4
Specific pricing-method indicator	28	0	11	17

Counts are code events. Coverage indicator C denotes carrier judgement and is not proof of affirmative coverage. Pricing indicator 00 or 99 means that no separate Part B pricing method or value was specified in the baseline file. Raw fields for every code are in Additional file 2.

3.2 Representation in the two annual files

Against that composition the primary result states itself quickly. Nineteen of 48 code events were represented by at least one national exact-code row in a public file covering services furnished during the effective year or the following year (39.6%, 95% Wilson interval 27.0 to 53.7%). Eleven were present in the file for the effective service year, eight were absent there and present in the file for the following service year, and 29 appeared in neither. By cohort year the pattern was 1 of 8 for 2021, 8 of 19 for 2022 and 10 of 21 for 2023 (Table 4; Fig. 2). Intervals are wide throughout and widest for the eight-code 2021 group.

Table 4. Representation by national exact-code row in the analytic cohort

	Overall	2021	2022	2023
Code events, N	48	8	19	21
National row in the file for service year Y, n	11	1	5	5
Absent in Y, national row in the file for $Y + 1$, n	8	0	3	5
No national row in either file, n	29	7	11	11
Represented, n/N	19/48	1/8	8/19	10/21
Represented, %	39.6	12.5	42.1	47.6
95% Wilson interval, %	27.0–53.7	2.2–47.1	23.1–63.7	28.3–67.6

The two files jointly cover services furnished from 1 April of the effective year through 31 December of the following year. A file for service year $Y + 1$ records services furnished in $Y + 1$; presence there does not show when the row was released. "No national row" refers to the specified public files as queried on 22 August 2026 and does not mean zero use.

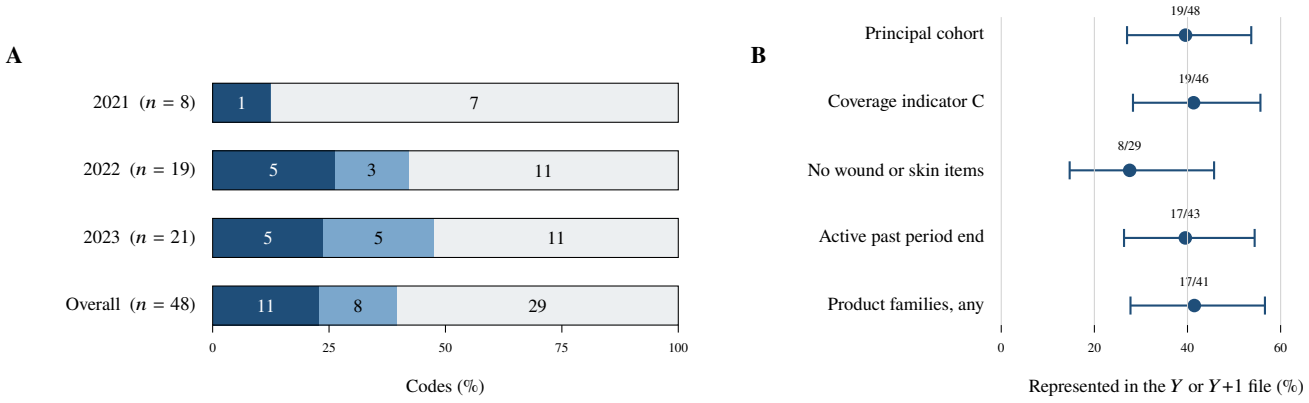


Figure 2. Representation states and sensitivity estimates. Panel A gives the mutually exclusive annual states for each April cohort and overall. Panel B gives the proportion represented in a file covering services furnished in Y or $Y + 1$, with 95% Wilson intervals; the family-level estimate counts a product family as represented when any member has a national row. An absent national row is an unresolved data state, not a zero-use observation.

Which dataset carried the row was clear-cut. Taking the earlier of the two files that held a row, eleven codes were represented in the physician file and eight in the equipment and supplies file, and no code appeared in more than one dataset in that year. Not one national outpatient hospital row appeared for any code in either service year. Physician representation was concentrated among wound and skin-substitute codes, while the equipment file carried the observed equipment, component, prosthetic and software records; the pattern describes where public rows were found and not where clinical activity occurred. All 288 cells are displayed in Additional file 9.

3.3 Representation by baseline subgroup

Baseline strata separate the cohort more sharply than the cohort-year split (Fig. 3, panel A). Wound and skin-substitute items were represented in 11 of 19 cases (57.9%, 36.3 to 76.9%), against 5 of 18 stand-alone items (27.8%, 12.5 to 50.9%) and 3 of 11 components, accessories or recurring supplies (27.3%, 9.7 to 56.6%). Neither of the two codes carrying coverage indicator S reached a public file, which is why restricting to indicator C raises the estimate from 19 of 48 to 19 of 46 without adding a single represented code. Baseline pricing splits the cohort along a similar line, with 16 of 28 codes with a specific pricing-method indicator were represented (57.1%, 39.1 to 73.5%) against 3 of 20 with indicator 00 or 99 (15.0%, 5.2 to 36.0%).

Four cautions govern how far any of that can be taken. Strata overlap heavily, because almost every wound or skin-substitute item carries a specific pricing method while the 2021 K-codes carry indicator 99 almost uniformly, so product class, pricing field and cohort year cannot be separated in

48 observations. Wilson intervals for the smaller strata run to thirty or forty percentage points. Codes within a product family are not independent. And every one of these fields is an administrative label recorded at code entry, not a measured property of the product or of clinical practice. Panel A of Fig. 3 is a description of the cohort, not a model of it.

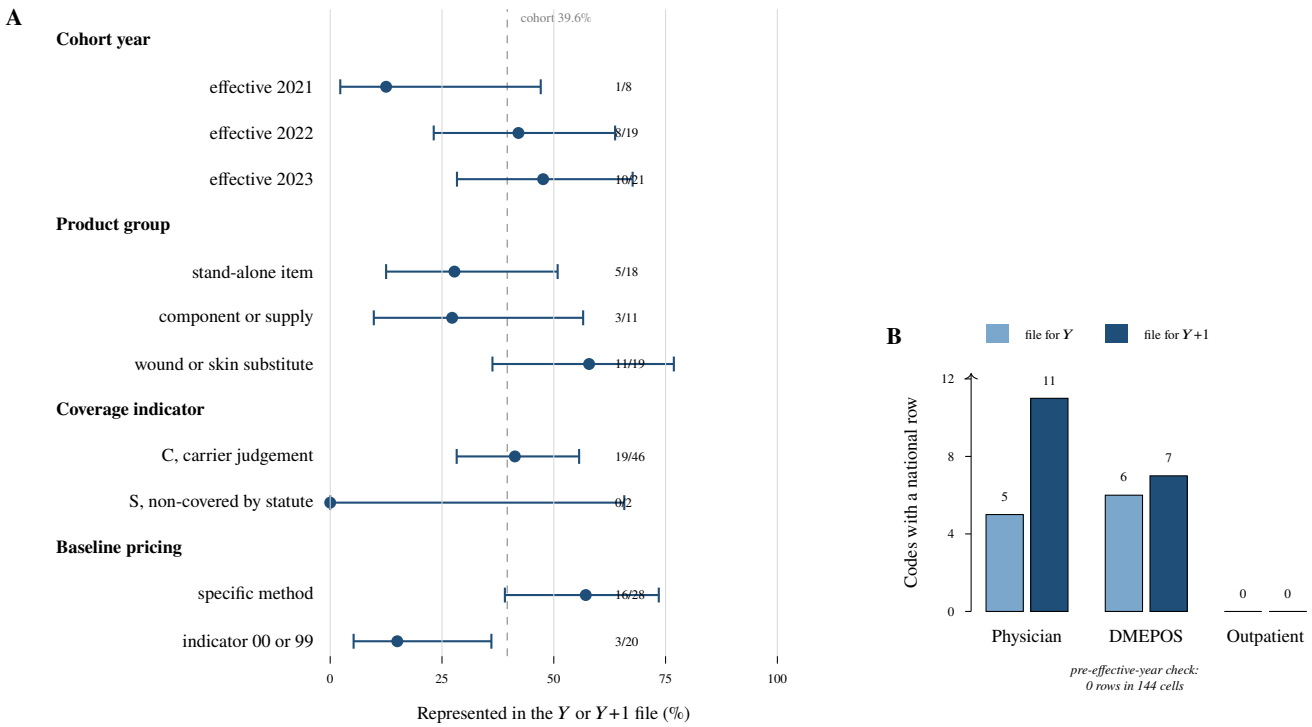


Figure 3. Representation by baseline stratum, and where the national rows sit. Panel A gives the proportion represented within each baseline stratum with 95% Wilson reference intervals, against the cohort value of 39.6% shown as a dashed line; strata overlap and are not adjusted for one another. Panel B counts, for each dataset and each of the two annual files, how many of the 48 codes carried a national exact-code row, with the pre-effective-year control beneath.

3.4 Where the national rows sit

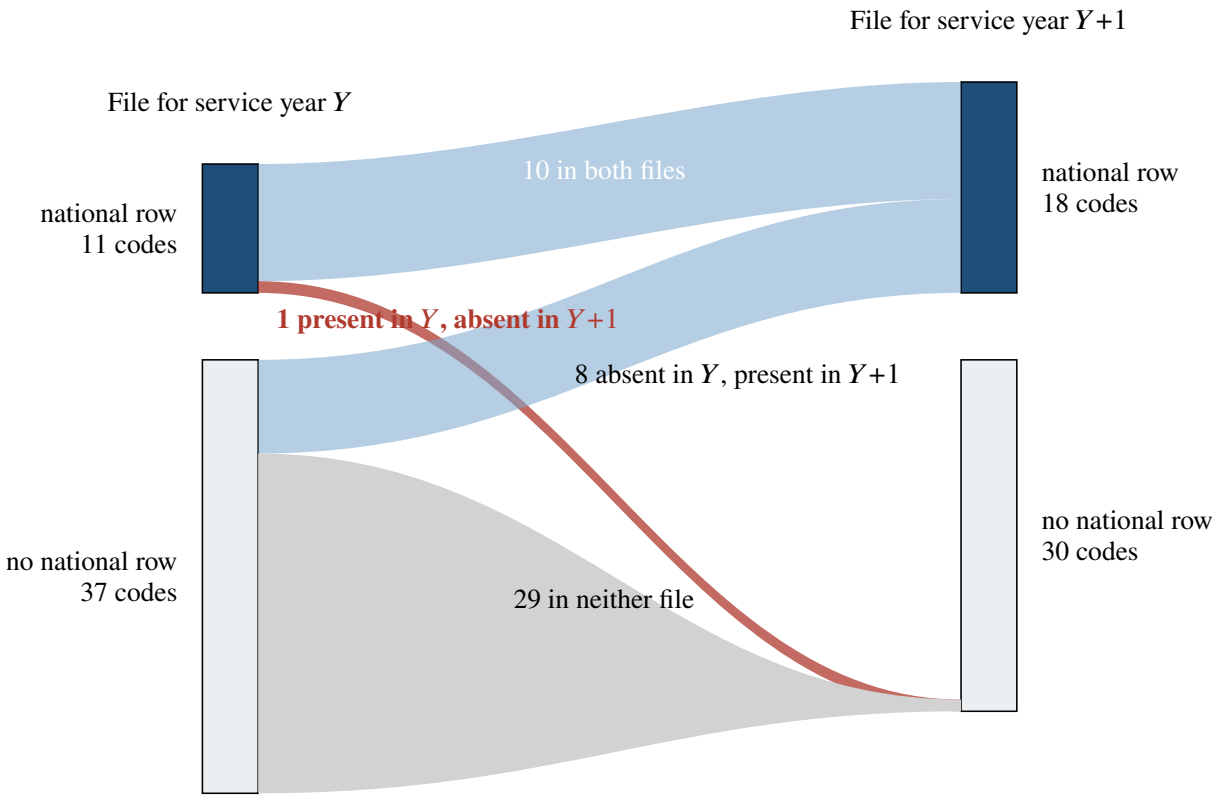
Twenty-nine of the 288 measurement cells held a national exact-code row, and their distribution across datasets is uneven in a way the cohort-level figure hides (Fig. 3, panel B). Physician files carried rows for 5 codes in the effective service year and for 11 in the following year; equipment files carried 6 and 7. The difference between the two annual files therefore sits almost entirely on the physician side, where six of the eight codes that appear only in the later file are wound or skin-substitute items, while the equipment picture is close to static: a code that appears in the equipment file at all tends to appear there already in the first year. Outpatient hospital files carried no row for any code in either service year, so all 96 outpatient cells were empty. Code eligibility for that file was not established from official documentation, so the outpatient query serves throughout as a secondary cross-domain check and not as a third equivalent channel.

For an informatics audit the practical consequence is that the two annual files, and the two datasets that contribute rows at all, are not interchangeable sources. Four cells held more than one national row: E2102 in the equipment file for service year 2023, and Q4265, Q4266 and Q4271 in the physician file for service year 2024. Each is a legitimate partition rather than a duplicate, by supplier rental status in the first case and by place of service in the other three. The binary rule counted each cell once and no beneficiary, service or supplier value was added across partitions, since the denominators of those fields differ between partitions of the same aggregate.

3.5 Persistence across annual files

Four states describe the cohort once both annual files are read together (Fig. 4). Ten codes carried a national row in both files, one in the effective-year file only, eight in the following-year file only, and 29 in neither. Of the eleven codes present in the file for their effective service year, ten were also present in the following-year file; A4341, a component of an intraurethral drainage system, is the exception, with a national equipment row for service year 2023 and none for 2024. Its family partner A4342 was never represented, so the family as a whole moved from one represented member to none.

Two consequences follow for how the endpoint may be modelled. Representation is not an absorbing condition, so treating the earlier file that shows a row as a failure time would suppress the observed reversal and describe the reporting process as monotone when the cohort shows that it is not. And the totals on the two sides of Fig. 4 differ, 11 against 18, only because the gains outnumber the single loss; a cohort in which they did not would show a falling count without any product leaving the market.



Representation does not accumulate monotonically: one code present in the earlier file was absent from the later one.

Figure 4. Movement of the 48 code events between the two annual files. Bar heights are proportional to the number of codes. The highlighted band is the single code represented in the file for the effective service year and absent from the file for the following year.

3.6 Sensitivity to cohort definition and code dependence

Estimates move with cohort composition but not with any single restriction (Table 5). Restriction to the 46 coverage-C codes gives 19 of 46; the 36-code decision-document subset gives 11 of 36, and 11 of 34 after removal of its two coverage-S codes. Removal of all 19 wound and skin-substitute items gives 8 of 29, and the same removal within coverage C gives 8 of 27. Exclusion of the five codes whose termination date falls on their service-period end leaves 17 of 43. Two boundary decisions were tested directly: dropping A4100, the one not-otherwise-specified code among the retentions, gives 19 of 47, and restoring S1091 to the denominator gives 19 of 49, since its archived responses hold no exact-code row in any of its six cells. At family level, 17 of 41 families had at least one represented member. Dropping the outpatient dataset from the outcome altogether, so that representation depends on the physician and equipment files alone, leaves the estimate unchanged at 19 of 48, because no outpatient cell held a national row in either service year. Leaving out one cohort year at a time produces 18 of 40 (45.0%), 11 of 29 (37.9%) and 9 of 27 (33.3%); those three rows and every restriction in Table 5 carry a uniquely identified line in Additional file 10. Intervals overlap across all of these, though the point estimates show that product composition moves the result by roughly twelve percentage points, which is more than any other restriction achieves.

We tested dependence between related codes at the level where it arises. Four of the six multi-member families are internally concordant, with either every member represented or none: both codes of the 2022 adjunctive continuous glucose monitoring pair were represented, and neither code of the two 2021 neurostimulation pairs or of the 2022 sleep-apnoea pair was. Two families split, the 2022 compression system at two of three members and the 2023 intraurethral drainage system at one of two. Counting each family once therefore changes the denominator by seven and the numerator by two, and moves the estimate from 39.6% to 41.5%. We report both family rules, because the any-member rule is asymmetric and can only raise the estimate: counting a family as represented when every member is gives 15 of 41 (36.6%, 23.6 to 51.9%), against 17 of 41 under the any-member rule, with 15 families fully represented, two partial and 24 with no represented member. The pattern is consistent with, but does not establish, dependence within a product system, and it is the reason the code-level intervals should not be read as if the 48 observations were independent.

Table 5. Sensitivity analyses

Analysis	Represented/*N*	Represented, %	95% Wilson interval, %
Analytic cohort of code events	19/48	39.6	27.0–53.7
Coverage indicator C	19/46	41.3	28.3–55.7
Decision-document subset	11/36	30.6	18.0–46.9
Decision-document subset, coverage C	11/34	32.4	19.1–49.2
Excluding wound and skin-substitute items	8/29	27.6	14.7–45.7
Coverage C and excluding those items	8/27	29.6	15.9–48.5
Codes usable beyond their service-period end	17/43	39.5	26.4–54.4
Excluding the not-otherwise-specified code A4100	19/47	40.4	27.6–54.7
Including S1091, the excluded S-series code	19/49	38.8	26.4–52.8
Dropping the outpatient dataset	19/48	39.6	27.0–53.7
Product families, any member represented	17/41	41.5	27.8–56.6
Product families, all members represented	15/41	36.6	23.6–51.9

The family-level denominator is smaller because codes assigned to one device system or express parent-and-component relation count once. Intervals are descriptive; they neither remove dependence between related codes nor convert the cohort into a random sample.

3.7 Pre-effective-year falsification check, lifecycle and data verification

None of the 144 pre-effective-year cells contained an exact-code row, whether national or state-level, which is what a correctly assembled cohort of newly established codes should produce. Every code remained usable throughout its assigned service period. K1028, K1029, K1031, K1032 and K1033 terminated on 31 December 2023, the final day of their period and, under the CMS record layout, the last day on which they could be used; their recorded successors became effective the following day. Eight 2021 K-codes terminated after their own service period, and thirteen later terminations carry locally documented successor mappings.

All 288 expected cells were present. Archived responses contained 392 exact-code rows, of which 359 were state-level rows retained for quality control and 33 were national rows. The 33 national rows fall in 29 cells, because four cells hold two legitimate national partitions each; the binary endpoint therefore counts 29 positive cells and never 33. Every non-empty response matched the requested code after case and whitespace normalisation.

Two codes complicate the linkage between an entry record and a later file, and we resolved both before reading the outcome. A4238 and E2102, the adjunctive continuous glucose monitoring pair, received descriptor revisions during follow-up; the exact code string is unchanged in both cases, so the link holds and the revision affects only the text of the descriptor, not the identity of the row. Thirteen later terminations carry successor mappings from an exact normalised descriptor match in the April 2025 file, with the successor effective on the day after termination; no analysis transfers an outcome from a predecessor to a successor.

Validation covered the full 49-code descriptor screen, that is the 48 analytic codes plus S1091, which is 49 codes times six cells and so 294 outcome cells against the 288 of the analytic cohort. It independently passed checks on those 294 cells, manifest byte counts, SHA-256 identity, exact-code values, decision-document routing and the absence of any national outpatient aggregate.

3.8 Exploratory contrast by baseline pricing indicator

One further comparison is reported descriptively and carries little weight. Among the 46 coverage-C codes, 16 of the 28 codes with a specific pricing-method indicator were represented (57.1%) against 3 of 18 with indicator 00 or 99 (16.7%). Removal of all wound and skin-substitute items reduces the specific-method group to nine codes, of which five were represented (55.6%), and leaves the 00 or 99 group at 3 of 18. A gap of forty percentage points in the first comparison and of thirty-nine in the second rests, after that single restriction, on nine codes. No test statistic is reported here, and none should be read into the counts: the groups are administrative categories in a purpose-built finite cohort, the cells are small, product families make the codes dependent, and pricing indicator is entangled with prefix, family and service type. Additional file 15 prints the same contrast under four cohort definitions with interval estimates for readers who want them.

4 Discussion

4.1 What the cohort shows

Fewer than half of the new supplied-product codes in this cohort, 19 of 48, were represented by a national exact-code row in three public Medicare fee-for-service files covering services furnished during the effective year or the year after it. Twenty-nine were not. Absence for twenty-nine codes is not an estimate of unused products; it measures how large the unresolved public-data state was for this cohort under the stated files, service years and retrieval date. Code effectiveness and representation in released data are separate administrative conditions, and satisfying the first does not establish the second. Read as an informatics result, the cohort measures a property of the released files and not of the products the codes name.

Annual detail adds information that a single cumulative percentage hides. Eleven codes were already represented in the file for their effective service year, eight only in the following-year file, and one code present in the earlier file was gone from the later one. A simple time-to-adoption reading cannot accommodate that reversal. What the pattern

fits is a reporting process in which clinical volume, claims domain, payment treatment, annual aggregation and release rules jointly determine whether a row becomes public, with the data unable to apportion the contribution of each.

Scale deserves a plain statement. If a researcher had drawn a sample of new product technologies from these public files for the 2021 to 2023 April cycles, the sample would have contained nineteen of the forty-eight codes that existed, and nothing in the released data would have signalled the absence of the other twenty-nine. Whether that matters for a particular study turns on whether the missing codes differ from the present ones in ways bearing on the research question, which the present data describe by baseline stratum without resolving.

4.2 The ceiling on inference

Annual files queried in 2026 record the service period they cover and nothing about their own release history. Presence of a code in the file for service year $Y + 1$ therefore means that services furnished in $Y + 1$ were represented when the file was queried; it does not mean that the code first became public in $Y + 1$, nor that it was absent from the file for Y, nor that any particular number of months elapsed before availability. Our reading of the seven-quarter window is correspondingly narrow, since it equalises the service period across cohorts by design rather than measuring an interval.

Readers who need release timing require the release history of each annual file, which the present sources do not carry. A single field recording when each row first entered a released version would convert the present binary outcome into a timing measure without any change of study design, and its absence is the reason a study of this kind cannot yet be run as a survival analysis. Until such a field exists, the honest object of measurement is a state at a documented retrieval date.

4.3 Why a national row may be missing

At least four mechanisms produce the same absence. Underlying service volume may fall below a public-release threshold. A claim may sit in a domain outside the three files, or the item may be bundled into another service. Use may occur in Medicare Advantage or outside Medicare altogether. Or no qualifying fee-for-service claim may exist. Public files collapse those states into one, which is why the conservative label used throughout is the variable actually observed rather than a rhetorical hedge.

Distinguishing among them requires different data, and each mechanism has a source that would settle it. Restricted claim-level files would separate low volume from absence of any claim and would locate a first paid service. Linkage to national and local coverage records would add the coverage event as a distinct field. Medicare Advantage encounter data would test the payer explanation. Supplier-level or provider-level files, used alongside national aggregates as Read and colleagues did for video capsule endoscopy, would show how much of the difference is a suppression artefact [5]. None of those sources is a substitute for the others, and a study that cannot obtain them should record the absence as unresolved and not resolve it by assumption.

The zero result in the outpatient hospital file deserves the same treatment. Packaging, setting, code suitability or reporting rules may each explain it, and channel eligibility for these particular codes was not established from the archived material. An absent outpatient row is therefore not evidence that hospital outpatient departments did not use the products. What can be said is narrower and still useful. For this class of codes, the outpatient file contributed nothing at all, so a study that queries only the physician and equipment files loses no information about them, as the sensitivity analysis confirms.

4.4 What the baseline pattern supports, and what it does not

Three baseline strata separate the cohort, and the separation is large enough to be worth naming even though the cohort cannot explain it. Wound and skin-substitute items were represented roughly twice as often as stand-alone items or components. Codes with a specific pricing method at entry were represented roughly four times as often as codes carrying indicator 00 or 99. Neither code with coverage indicator S was represented at all.

Reading any of that as a causal statement about payment would mistake the design, and not only because the cells are small. Product class, pricing field, code prefix, service type and cohort year move together in this cohort by construction: every 2021 code is a K-code with indicator 99 and none is a wound product, while the wound and skin-substitute items that arrive in 2022 and 2023 carry specific pricing methods almost throughout. A regression on 48 observations with that structure could not separate those attributes stably, and its coefficients would carry no credible causal reading. A defensible reading stays descriptive and conditional. Within this cohort, the represented codes were disproportionately wound and skin-substitute products with a specific pricing method at entry, and the codes that were not were disproportionately stand-alone devices and their supplies entering in the earliest cycle. Whether the pricing field marks an item expected to generate volume, or whether volume and pricing both follow from something else, awaits a design with a larger cohort and access to claim-level data.

4.5 Asymmetry between the two annual files

The difference between the two annual files is not distributed evenly across the datasets, and the asymmetry is one of the more informative details in the cohort. Physician files went from five represented codes to eleven, equipment files from six to seven. Six of the eight codes that appear only in the later file are wound or skin-substitute items billed on the physician side. Equipment codes behave differently, in that those reaching the file at all tend to be there in the first year, while those absent in the first year mostly never appear.

Gradual accumulation of low-volume practitioner billing and concentration of supplier billing among few suppliers are both compatible with the pattern, and aggregate data cannot distinguish them. What follows regardless is operational: a researcher querying only the file for the effective service year would have found eleven codes where two annual files find nineteen, and the loss would fall mainly on one product class in one claims domain.

4.6 Relation to adjacent research

No design that begins from a technology already present in the files could reach either reading, and that is where most of the adjacent work begins. Read and colleagues could describe the diffusion of video capsule endoscopy over seventeen years because a stable dedicated procedure code was already represented across many annual files [5]; Tang and colleagues could describe early growth in remote patient monitoring for the same reason [4]. The present cohort sits one step earlier and asks whether a newly effective product code reaches released national data at all during a fixed initial service period. Once it does and volume accumulates, longitudinal measures become available; before it does, a researcher has no public numerator and cannot separate low use from administrative absence. Both approaches complement each other and do not compete.

Table 1 places the finding against the six strands set out in the Background, and the placement carries an implication for two of them. Designs that measure utilisation from released claims, and designs that evaluate a payment reform through utilisation in those claims, both draw their samples from the represented side of a gate whose size is now estimated for one class of codes. A payment-policy evaluation that finds no utilisation response for a newly coded technology cannot distinguish, from public data alone, between a technology that was not used and a code that was not released. Research on device coding has largely emphasised insufficient specificity, and unique device identifier projects aim to carry item identity from the point of care into records and claims [24, 25, 26, 27]. A second and simpler obstacle appears here even after a dedicated HCPCS code exists, in that the public aggregate may still be absent. Specificity of a code and representation of that code in released data are separate conditions, and meeting one does not establish the other. Soril and colleagues showed how far technology questions can be translated into code-based queries against administrative data while keeping clinical adjudication inside the method [50]; the same discipline applies to the reporting gate, which is a property of the data rather than of the technology.

4.7 Consequences for study design and policy evaluation

An event ledger, not a single date labelled access, is what early technology studies need. At minimum such a ledger keeps the code identity, the effective date, the raw benefit and pricing fields, any national or local coverage source, each queried claims domain, the annual row state and the reason a missing value remains unresolved. Restricted claims that later show a paid service add a date without rewriting the earlier coding event, and a disappearance from a later public file stays in the ledger as a reversal instead of being smoothed away.

Operationally the rule is small. A researcher evaluating a newly coded technology can begin with six binary fields, one per dataset in each of the two service years, and can settle the natural key of each source before any volume is analysed, since place-of-service and rental partitions must not be added naively. Missing public rows stay unknown until a stronger source resolves them, and Table 5 is the template for asking whether a result depends on product class, benefit treatment or linked-code families.

Reporting deserves the same discipline, and it is here that the measurement question becomes a matter of data governance. A study built on public aggregates should state its retrieval date, because the files are revised; should state which datasets were queried and which were not, because a zero in an unqueried domain is not a zero; and should report the denominator of codes that were sought as well as the numerator of codes that were found. None of that is onerous, and its absence is what makes a claims-first sample impossible to audit from the outside.

Policy evaluation benefits from the same discipline. Coding, coverage and payment run as separate processes under CMS's own account [6], and a reform that shortens an authorisation-to-coverage interval need not shorten code establishment, provider acquisition, claims accumulation or public release. An evaluation that assigns one date to the whole pathway can credit or blame the wrong institution. Cross-country work shows that the arrangement of these institutions varies, so the specific numbers here are Medicare-specific while the requirement to separate the events is not [9].

4.8 Strengths and limitations

What supports the analysis is mostly a matter of where it starts. A complete set of April additions entered the screen, so the design does not start from products already found in claims, which is what makes the unrepresented group observable at all. Every cohort year shares the same 1 April effective date and seven-quarter service period, avoiding the unequal exposure that mixed quarterly starts would produce. Exact code matching, a pre-effective-year control, lifecycle adjudication, per-response hash records and a full cell matrix make the result auditable at code level.

Limitations set a firm ceiling. The descriptor screen was conducted at code level by one author and establishes no regulatory status for any wound or skin-substitute product, which is why all such codes stay together in the main analysis and are removed together in sensitivity analysis. Three April releases and 48 codes are a small base, the 2021 group especially so, and code families create dependence even though the family-level estimate was similar. Public geography-and-service files are annual, aggregate and limited to specified Original Medicare fee-for-service domains; they carry no first claim date, no individual pathway, no denial, no clinical indication and no outcome. Suppression prevents separation of low volume from absence of use, Medicare Advantage and other payers are absent, and different definitions of service units and provider counts across files are the reason no volumes were combined.

Baseline coverage and pricing indicators are administrative labels rather than adjudicated coverage policies or transaction prices, and their overlap with product class rules out a credible causal analysis in a cohort of this size. Wilson intervals neither remove code dependence nor convert the cohort into a random sample of future coding actions. Data retrieved later may differ if CMS revises a public release; the archived hashes state which version produced this result but cannot freeze the external repository. Nothing here supports

generalisation beyond April cycles, beyond supplied-product codes or beyond three datasets, and a cohort drawn from January, July or October additions would carry service periods of different length and would need separate analysis, not pooling with this one.

5 Conclusions

Among 48 newly established HCPCS supplied-product codes effective on 1 April in 2021 to 2023, nineteen were represented by a national exact-code row in the queried public Medicare fee-for-service files covering services furnished in the effective year or the following year. Eleven appeared in the file for the effective service year and eight only in the next; twenty-nine appeared in neither, and one code present in the earlier file was absent from the later one. No result licenses calling those twenty-nine codes unused, uncovered or inaccessible, and none of the estimates here describes when CMS released anything.

What the finding changes is the standing of a common design. A study that assembles its sample from codes already present in released claims is restricted to code events represented in those files. In this cohort 19 of 48 code events met that definition, and the remaining 29 would not have entered such a sample at all. For health services research that treats these files as a population register, the representation state is part of the study design and not a footnote to it. Whether the unrepresented part matters depends on the question being asked, but its size can no longer be assumed to be small, and a paper that samples this way should say so rather than leave the reader to infer it.

Applying the rule takes little effort. Record code effectiveness, coverage and pricing fields, claims domain, annual row state and product-family dependence as separate fields; permit representation to reverse between annual files; establish the natural key of a source before adding any volume across partitions; state the retrieval date and the datasets queried; report the denominator of codes sought alongside the numerator found; and leave an absent public row unresolved rather than reading it as zero. Followed consistently, that ledger replaces an impression of slowness with a measurement rule another researcher can reproduce, and it keeps the architecture of the data from being mistaken for the behaviour of patients and clinicians.

One change to how the public files are published would retire the central limitation of this design. Versioned releases with stable release identifiers, a change log and row-level provenance recording the first and last release in which a row occurs would turn representation from a state observed at one retrieval date into a measurable interval. A single mutable first-appearance field would not, because public aggregates are revised and rows are sometimes withdrawn. Absent that, the measurement rule above is what a researcher can defend. Governance of public health data and the research built on it meet here in one practical point: the twenty-nine unrepresented codes remain what the data actually say they are, an unresolved state and not an empty one.

Abbreviations

CMS: Centers for Medicare & Medicaid Services; DMEPOS: durable medical equipment, prosthetics, orthotics and supplies, used here as the short name for the CMS Medicare Durable Medical Equipment, Devices & Supplies public-use file; FDA: US Food and Drug Administration; HCPCS: Healthcare Common Procedure Coding System; RECORD: REporting of studies Conducted using Observational Routinely-collected health Data; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology.

Supplementary information

Each additional file is listed below by file name, format, title and description, and each is cited in sequence in the text above. The delivered data files carry column names matching the endpoint as defined in the Methods and an unambiguous cohort marker; the upstream analysis output ships unaltered inside Additional file 16, and Additional files 12 and 13 record every rename and every changed value between the two.

Additional file 1 `Additional_file_1_screening_ledger.csv` (CSV). Descriptor screen. All 103 April additions with product class, screen disposition, analytic-cohort membership and decision-document routing.

Additional file 2 `Additional_file_2_code_ledger.csv` (CSV). Analytic cohort. The 48 code events with raw HCPCS fields, family assignment, annual representation states, falsification check and lifecycle adjudication.

Additional file 3 `Additional_file_3_measurement_cells.csv` (CSV). Measurement cells. All 288 code by service year by dataset cells with row counts and the national-row indicator.

Additional file 4 `Additional_file_4_falsification_cells.csv` (CSV). Pre-effective-year falsification check. The 144 control cells.

Additional file 5 `Additional_file_5_code_lifecycle.csv` (CSV). Code lifecycle. Termination date, relation to the service period and successor mapping for every code.

Additional file 6 `Additional_file_6_code_family_map.csv` (CSV). Product families. The 48 analytic codes in 41 families, with the basis of each assignment.

Additional file 7 `Additional_file_7_source_manifest.csv` (CSV). Retrieval manifest. Dataset identifier, request URL, retrieval date, byte count and SHA-256 hash for each main query.

Additional file 8 `Additional_file_8_falsification_source_manifest.csv` (CSV). The same for the 144 control queries.

Additional file 9 `Additional_file_9_analysis_summary.csv` (CSV). Machine-readable summary of the principal analysis.

Additional file 10 `Additional_file_10_sensitivity_results.csv` (CSV). Canonical sensitivity table. One uniquely identified row for each of the 15 analyses reported in this article, the reference estimate included.

Additional file 11 `Additional_file_11_decision_document_crosscheck.csv` (CSV). Decision-document routing check for the 49 screened codes.

Additional file 12 `Additional_file_12_column_mapping.csv` (CSV). Every column renamed between the upstream analysis output and the delivered files.

Additional file 13 `Additional_file_13_corrections.csv` (CSV). Every value changed against the upstream output, with the type of change and the reason.

Additional file 14 `Additional_file_14_exact_code_matrix.pdf` (PDF). Exact-code matrix for all 288 cells.

Additional file 15 `Additional_file_15_supplementary_information.pdf` (PDF). Supplementary information: data dictionary, dataset identifiers and request template, full cohort and lifecycle listings, every sensitivity estimate, the pricing contrast, the exact-code matrix and the reporting map.

Additional file 16 `Additional_file_16_reproducer.zip` (ZIP). Reproducer. The data-build and figure scripts together with the unaltered upstream analysis output they read. Unpacked into an empty directory, the documented command regenerates Additional files 1 to 13 and the four figures. It does not reproduce the CMS retrieval, does not rerun the upstream analysis, and does not rebuild Additional file 15, which needs a LaTeX toolchain.

Acknowledgements

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable. The study uses public aggregate administrative data and contains no individual-level or identifiable record.

Consent for publication

Not applicable.

Availability of data and materials

All source data are public CMS records, retrievable by anyone from the URLs listed in the supplementary material. The datasets supporting the conclusions of this article are included within the article and its additional files: the derived ledgers, the complete outcome and falsification matrices, the canonical sensitivity table, the retrieval manifests giving the request URL, retrieval date, byte count and SHA-256 hash of each of the 438 archived responses, and the deterministic data-build and figure code with the frozen inputs it reads. Archived CMS response payloads themselves are held by the author and are not included here; they will be deposited in a public repository together with the upstream analysis code and a catalogue snapshot, and that deposit will be cited by its persistent identifier in this statement and in the reference list.

Competing interests

No competing interests are declared.

Funding

No external funding was received for this study.

Authors' contributions

AP conceived the study, defined the research question, developed the measurement design, and is responsible for the final methodology, analysis, interpretation and manuscript.

Author information

ORCID: 0009-0006-0059-1420.

References

[1] Han M, Jing Y, Hammadeh Z, Lee EN, Chen JX, Cronin C, et al. Medical Device Industry Payments and Utilization of Corresponding Devices. JAMA Network Open. 2026;9(7):e2623967. https://doi.org/10.1001/jamanetworkopen.2026.23967.

[2] Hotz A, Sprecher E, Bastianelli L, Rodean J, Stringfellow I, Barkoudah E, et al. Categorization of a Universal Coding System to Distinguish Use of Durable Medical Equipment and Supplies in Pediatric Patients. JAMA Network Open. 2023;6(10):e2339449. https://doi.org/10.1001/jamanetworkopen.2023.39449.

[3] Holtzman JN, Wadhera RK, Choi E, Zhao T, Secemsky EA, Fraiche AM, et al. Trends in Utilization and Spending on Remote Monitoring of Pacemakers and Implantable Cardioverter-Defibrillators Among Medicare Beneficiaries. Heart Rhythm. 2020;17(11):1917–1921. https://doi.org/10.1016/j.hrthm.2020.05.044.

[4] Tang M, Nakamoto CH, Stern AD, Mehrotra A. Trends in Remote Patient Monitoring Use in Traditional Medicare. JAMA Internal Medicine. 2022;182(9):1005–1006. https://doi.org/10.1001/jamainternmed.2022.3043.

[5] Read AJ, Rice MD, Baker JR, Waljee AK, Saini SD. Diffusion of an Innovation: Growth in Video Capsule Endoscopy in the U.S. Medicare Population From 2003 to 2019. BMC Health Services Research. 2022;22:425. https://doi.org/10.1186/s12913-022-07780-2.

[6] Centers for Medicare & Medicaid Services. HCPCS Level II Coding Procedures. Centers for Medicare & Medicaid Services; 2026. Available from: https://www.cms.gov/medicare/coding-billing/healthcare-common-procedure-system/level-ii-coding-process.

[7] Ciani O, Wilcher B, van Giessen A, Taylor RS. Linking the Regulatory and Reimbursement Processes for Medical Devices: The Need for Integrated Assessments. Health Economics. 2017;26(S1):13–29. https://doi.org/10.1002/hec.3479.

[8] Chambers JD, May KE, Neumann PJ. Medicare Covers the Majority of FDA-Approved Devices and Part B Drugs, but Restrictions and Discrepancies Remain. Health Affairs. 2013;32(6):1109–1115. https://doi.org/10.1377/hlthaff.2012.1073.

[9] Williams MM, Smith NR, Uyl-de Groot CA, den Uil CA, Ross JS, Mohamed MO, et al. Variations in the Medical Device Authorization and Reimbursement Landscape: A Case Study of 2 Cardiovascular Devices Across 4 Countries. Circulation: Cardiovascular Quality and Outcomes. 2025;18(4):e011636. https://doi.org/10.1161/CIRCOUTCOMES.124.011636.

[10] Sexton ZA, Perl JR, Saul HR, Trotsyuk AA, Pietzsch JB, Ruggles SW, et al. Time From Authorization by the US Food and Drug Administration to Medicare Coverage for Novel Technologies. JAMA Health Forum. 2023;4(8):e232260. https://doi.org/10.1001/jamahealthforum.2023.2260.

[11] Zinn AM, Allen JC Jr, Hacker CS. Median Approval Times for Class III Medical Devices Have Been Well Above Statutory Deadlines Set for FDA and CMS. Health Affairs. 2012;31(10):2304–2313. https://doi.org/10.1377/hlthaff.2010.1198.

[12] Roginiel AC, Dhruva SS, Ross JS. Evidence Supporting FDA Approval and CMS National Coverage Determinations for Novel Medical Products, 2005 Through 2016: A Cross-Sectional Study. Medicine. 2018;97(40):e12715. https://doi.org/10.1097/MD.0000000000012715.

[13] Chambers JD, Chenoweth M, Thorat T, Neumann PJ. Private Payers Disagree With Medicare Over Medical Device Coverage About Half the Time. Health Affairs. 2015;34(8):1376–1382. https://doi.org/10.1377/hlthaff.2015.0133.

[14] Moneer O, Mooghali M, Moosa K, Ramachandran R, Ross JS, Dhruva SS. Medicare Local Coverage Determinations: Evidence Quality Is Stronger for Covered Indications. Health Affairs. 2024;43(12):1712–1718. https://doi.org/10.1377/hlthaff.2024.00182.

[15] Kane NM, Manoukian PD. The Effect of the Medicare Prospective Payment System on the Adoption of New Technology: The Case of Cochlear Implants. New England Journal of Medicine. 1989;321(20):1378–1383. https://doi.org/10.1056/NEJM198911163212006.

[16] Shih C, Berliner E. Diffusion of New Technology and Payment Policies: Coronary Stents. Health Affairs. 2008;27(6):1566–1576. https://doi.org/10.1377/hlthaff.27.6.1566.

[17] Yang CM. The Impact of Global Budget on the Diffusion of Innovations: The Example of Positron Emission Tomography in Taiwan. BMC Health Services Research. 2018;18:905. https://doi.org/10.1186/s12913-018-3731-4.

[18] Morey JR, Katana M, Li L, Hohmann S, Weber E, Ferket B. Adoption and Trends in the Medicare New Technology Add-On Payment Program. Journal of General Internal Medicine. 2021;36(7):2174–2176. Published online in 2020; assigned to the 2021 print issue. https://doi.org/10.1007/s11606-020-05930-y.

[19] Judson TJ, Dhruva SS, Redberg RF. Evaluation of Technologies Approved for Supplemental Payments in the United States. BMJ. 2019;365:l2190. https://doi.org/10.1136/bmj.l2190.

[20] Moneer O, Johnston JL, Rathi VK, Ross JS, Dhruva SS. Medical Devices Applying for Outpatient Medicare Supplemental Payments. JAMA Health Forum. 2024;5(11):e244016. https://doi.org/10.1001/jamahealthforum.2024.4016.

[21] Xing Q, Chen T, Hu H, He D, Wang W, Yin Y, et al. The Impact of Deep Brain Stimulation Payment Adjustments on Healthcare Utilization in Shanghai, China: A Regression Discontinuity Analysis of Health Insurance Claims. BMC Health Services Research. 2025;25:1080. https://doi.org/10.1186/s12913-025-13301-8.

[22] Moscovitch B, Rising JP. Medical Device Identification in Claims Data. JAMA. 2017;318(19):1936–1937. https://doi.org/10.1001/jama.2017.15146.

[23] Drozda JP Jr, Roach J, Forsyth T, Helmering P, Dummitt B, Tcheng JE. Constructing the Informatics and Information Technology Foundations of a Medical Device Evaluation System: A Report From the FDA Unique Device Identifier Demonstration. Journal of the American Medical Informatics Association. 2018;25(2):111–120. https://doi.org/10.1093/jamia/ocx041.

[24] Zerhouni YA, Krupka DC, Graham J, Landman A, Li A, Bhatt DL, et al. UDI2Claims: Planning a Pilot Project to Transmit Identifiers for Implanted Devices to the Insurance Claim. Journal of Patient Safety. 2021;17(8):e708–e715. Published online in 2018; assigned to the 2021 print issue. https://doi.org/10.1097/PTS.0000000000000543.

[25] Krupka DC, Graham J, Wilson NA, Li A, Landman AB, Bhatt DL, et al. Transmitting Device Identifiers of Implants From the Point of Care to Insurers: A Demonstration Project. Journal of Patient Safety. 2021;17(3):223–230. https://doi.org/10.1097/PTS.0000000000000828.

[26] Drozda JP Jr, Graham J, Muhlestein JB, Tcheng JE, Roach J, Forsyth T, et al. Multi-Institutional Distributed Data Networks for Real-World Evidence About Medical Devices: Building Unique Device Identifiers Into Longitudinal Data (BUILD). JAMIA Open. 2022;5(2):ooac035. https://doi.org/10.1093/jamiaopen/ooac035.

[27] Dhruva SS, Ridgeway JL, Ross JS, Drozda JP Jr, Wilson NA. Exploring Unique Device Identifier Implementation and Use for Real-World Evidence: A Mixed-Methods Study With NESTcc Health System Network Collaborators. BMJ Surgery, Interventions, & Health Technologies. 2023;5(1):e000167. https://doi.org/10.1136/bmjsit-2022-000167.

[28] Wang X, Panickan VA, Cai T, Xiong X, Cho K, Cai T, et al. Endovascular Aneurysm Repair Devices as a Use Case for Postmarketing Surveillance of Medical Devices. JAMA Internal Medicine. 2023;183(10):1090–1097. https://doi.org/10.1001/jamainternmed.2023.3562.

[29] Choi S, Kim JK, Lee J, Choi SJ, Lee YK. Limitations of NHIC Claim Code-Based Surveillance and the Necessity of UDI Implementation in Korea. Scientific Reports. 2024;14(1):21014. https://doi.org/10.1038/s41598-024-72063-1.

[30] Tarricone R, Torbica A, Drummond M. Challenges in the Assessment of Medical Devices: The MedtecHTA Project. Health Economics. 2017;26(S1):5–12. https://doi.org/10.1002/hec.3469.

[31] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. The Lancet. 2007;370(9596):1453–1457. https://doi.org/10.1016/S0140-6736(07)61602-X.

[32] Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLOS Medicine. 2015;12(10):e1001885. https://doi.org/10.1371/journal.pmed.1001885.

[33] Centers for Medicare & Medicaid Services. April 2021 Alpha-Numeric HCPCS File. Centers for Medicare & Medicaid Services; 2021. Available from: https://www.cms.gov/files/zip/april-2021-alpha-numeric-hcpcs-file.zip.

[34] Centers for Medicare & Medicaid Services. April 2022 Alpha-Numeric HCPCS File. Centers for Medicare & Medicaid Services; 2022. Available from: https://www.cms.gov/files/zip/april-2022-alpha-numeric-hcpcs-file.zip.

[35] Centers for Medicare & Medicaid Services. April 2023 Alpha-Numeric HCPCS File. Centers for Medicare & Medicaid Services; 2023. Available from: https://www.cms.gov/files/zip/april-2023-alpha-numeric-hcpcs-file.zip.

[36] Centers for Medicare & Medicaid Services. Healthcare Common Procedure Coding System (HCPCS) Level II Final Coding Decisions: Second Biannual, 2020 Durable Medical Equipment and Accessories Coding Cycle. Centers for Medicare & Medicaid Services; 2021. Available from: https://www.cms.gov/files/document/2020-hcpcs-application-summary-bi-annual-2-2020-durable-medical-equipment-dme-and-accessories.pdf.

[37] Centers for Medicare & Medicaid Services. Healthcare Common Procedure Coding System (HCPCS) Level II Final Coding, Benefit Category and Payment Determinations: Second Biannual, 2021 Non-Drug and Non-Biological Items and Services Coding Cycle. Centers for Medicare & Medicaid Services; 2022. Available from: https://www.cms.gov/files/document/2021-hcpcs-application-summary-biannual-2-2021-non-drug-and-non-biological-items-and-services.pdf.

[38] Centers for Medicare & Medicaid Services. Healthcare Common Procedure Coding System (HCPCS) Level II Final Coding, Benefit Category and Payment Determinations: Second Biannual (B2), 2022 HCPCS Coding Cycle. Centers for Medicare & Medicaid Services; 2023. Available from: https://www.cms.gov/files/document/2022-hcpcs-application-summary-biannual-2-2022-non-drug-and-non-biological-items-and-services.pdf.

[39] Centers for Medicare & Medicaid Services. Note Regarding Coverage and Payment Indicators for Codes in CMS' 2020 HCPCS Update and DMEPOS Fee Schedule Files. Centers for Medicare & Medicaid Services; 2020. Accompanies the annual alpha-numeric HCPCS file. States that absent coverage or payment indicators for new codes indicate neither coverage nor non-coverage. Available from: https://www.cms.gov/files/document/2020-hcpcs-alpha-numeric-hcpcs-files-disclaimer.txt.

[40] Centers for Medicare & Medicaid Services. Medicare Physician & Other Practitioners – by Geography and Service. Centers for Medicare & Medicaid Services; 2026. Official CMS public-use dataset; page last modified May 21, 2026; temporal coverage through 2024 in the CMS data catalog. Available from: https://data.cms.gov/provider-summary-by-type-of-service/medicare-physician-other-practitioners/medicare-physician-other-practitioners-by-geography-and-service.

[41] Centers for Medicare & Medicaid Services. Medicare Durable Medical Equipment, Devices & Supplies – by Geography and Service. Centers for Medicare & Medicaid Services; 2026. Official CMS public-use dataset; page last modified July 23, 2026; temporal coverage through 2024 in the CMS data catalog. Available from: https://data.cms.gov/provider-summary-by-type-of-service/medicare-durable-medical-equipment-devices-supplies/medicare-durable-medical-equipment-devices-supplies-by-geography-and-service.

[42] Centers for Medicare & Medicaid Services. Medicare Outpatient Hospitals – by Geography and Service. Centers for Medicare & Medicaid Services; 2026. Official CMS public-use dataset; page last modified June 25, 2026; temporal coverage through 2024 in the CMS data catalog. Available from: https://data.cms.gov/provider-summary-by-type-of-service/medicare-outpatient-hospitals/medicare-outpatient-hospitals-by-geography-and-service.

[43] Centers for Medicare & Medicaid Services. Medicare Fee-For-Service Provider Utilization & Payment Data, Physician and Other Practitioners Dataset: A Methodological Overview. Centers for Medicare & Medicaid Services; 2025. Prepared by the CMS Office of Enterprise Data and Analytics. Available from: https://data.cms.gov/sites/default/files/2025-04/MUP_PHY_RY25_20250312_Methodology_508.pdf.

[44] Centers for Medicare & Medicaid Services. Medicare Durable Medical Equipment, Devices & Supplies Methodology. Centers for Medicare & Medicaid Services; 2025. Prepared by the CMS Office of Enterprise Data and Analytics. Available from: https://data.cms.gov/sites/default/files/2025-06/a837dcf7-3d06-4c1f-8d4a-82ee737b7932/MUP_DME_RY25_Methodology_20250617_508.pdf.

[45] Centers for Medicare & Medicaid Services. Medicare Outpatient Hospitals Dataset: A Methodological Overview. Centers for Medicare & Medicaid Services; 2026. Prepared by the CMS Office of Enterprise Data and Analytics. Available from: https://data.cms.gov/sites/default/files/2026-06/1a3abf36-bf10-46ae-a26a-d536d145ffbc/MUP_OUT_RY25_20260624_Methodology-508.pdf.

[46] Centers for Medicare & Medicaid Services. April 2025 Alpha-Numeric HCPCS File and Record Layout. Centers for Medicare & Medicaid Services; 2025. Available from: https://www.cms.gov/files/zip/april-2025-alpha-numeric-hcpcs-file.zip.

[47] Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. Journal of the American Statistical Association. 1927;22(158):209–212. https://doi.org/10.1080/01621459.1927.10502953.

[48] Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical Tests, *P* Values, Confidence Intervals, and Power: A Guide to Misinterpretations. European Journal of Epidemiology. 2016;31(4):337–350. https://doi.org/10.1007/s10654-016-0149-3.

[49] Wasserstein RL, Lazar NA. The ASA Statement on *p*-Values: Context, Process, and Purpose. The American Statistician. 2016;70(2):129–133. https://doi.org/10.1080/00031305.2016.1154108.

[50] Soril LJJ, Seixas BV, Mitton C, Bryan S, Clement FM. Moving Low Value Care Lists Into Action: Prioritizing Candidate Health Technologies for Reassessment Using Administrative Data. BMC Health Services Research. 2018;18:640. https://doi.org/10.1186/s12913-018-3459-1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- HCPCSsupplementaryinformation.pdf
- REPOSITORYdataandcode1.zip