# Supplementary information

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

## How to read this document

Sixteen additional files accompany the article. Sections 1 to 9 of this document describe what each file contains, define every column in the two central data files, list the exact CMS datasets that were queried, and print in full the material that the article reports only in summary. Nothing here is new evidence. Every number was produced by the deterministic analysis whose outputs are shipped alongside, and every table in this document is generated from those outputs by `analysis/build_supplement.py` without manual entry.

Three terms are used throughout in the sense fixed by the article. A *code event* is a five-character HCPCS Level II code newly established in an April update. A *measurement cell* is one code, one service year and one dataset. *Represented* means that the queried file carried at least one national row with the exact code when the file was retrieved on 22 August 2026; it is a property of a released file and says nothing about when that file, or that row, was first published. An empty result is an unresolved state and never a count of zero use.

## 1 Inventory of the additional files

Sixteen additional files accompany the article. Table S1 names each one, states what it holds and how large it is, and gives the leading characters of its SHA-256 hash so that a reader can confirm which version was received.

**Table S1.** Files delivered with the article. The digest column carries the first sixteen hexadecimal characters of the SHA-256 hash; the full hashes are in the `INVENTORY.csv` delivered with the upload package.

| File | Name | Contents | Size | Digest |
|---|---|---|---|---|
| 1 | `Additional_file_1_screening_ledger.csv` | Descriptor screen, all 103 April additions | 26 957 B | `fdfe98ea7233ca85` |
| 2 | `Additional_file_2_code_ledger.csv` | Analytic cohort, 48 code events | 20 139 B | `22e65ce663575e95` |
| 3 | `Additional_file_3_measurement_cells.csv` | Measurement cells, all 288 | 44 207 B | `04fa42180a4d988c` |
| 4 | `Additional_file_4_falsification_cells.csv` | Pre-effective-year falsification check, 144 cells | 13 052 B | `3ac0dcde231f0006` |
| 5 | `Additional_file_5_code_lifecycle.csv` | Code lifecycle and successor mapping | 9 432 B | `25f671cfaf03a847` |
| 6 | `Additional_file_6_code_family_map.csv` | Product-family assignment | 2 585 B | `2c5bc8d5ef23b9c3` |
| 7 | `Additional_file_7_source_manifest.csv` | Retrieval manifest, main queries | 87 130 B | `a10d3018ef95a3a9` |
| 8 | `Additional_file_8_falsification_source_manifest.csv` | Retrieval manifest, falsification queries | 48 212 B | `1b878297d6e53e27` |
| 9 | `Additional_file_9_analysis_summary.csv` | Machine-readable analysis summary | 3 156 B | `7523ea3012b920ed` |
| 10 | `Additional_file_10_sensitivity_results.csv` | Canonical sensitivity table, every reported analysis | 2 021 B | `180ce8d95c9427c6` |
| 11 | `Additional_file_11_decision_document_crosscheck.csv` | Decision-document routing check | 11 538 B | `05d3c68b0ca7d6ca` |
| 12 | `Additional_file_12_column_mapping.csv` | Column mapping against the upstream output | 6 375 B | `7884dd8c9d72c4a2` |
| 13 | `Additional_file_13_corrections.csv` | Log of every changed value | 5 602 B | `fa7a4d8c85c17408` |
| 14 | `Additional_file_14_exact_code_matrix.pdf` | Exact-code matrix, all 288 cells | 27 360 B | `6d5e2b6ec49b0eae` |
| 15 | `Additional_file_15_supplementary_information.pdf` | This document | see INVENTORY | see INVENTORY |
| 16 | `Additional_file_16_reproducer.zip` | Reproducer: scripts and the inputs they read | see INVENTORY | see INVENTORY |

Additional file 7 records, for each of the 438 archived responses, the exact request URL, the retrieval date, the byte count and the SHA-256 hash of the stored payload. A reader who wants to confirm that the analysis used the data it claims to have used should start there.

## 2 Column mapping and corrections

Two tables document the difference between the upstream analysis output and the delivered submission data set. Table S10, printed with the other long listings at the end of this document, lists every column that was renamed, and Table S2 below every value that was changed. Nothing else differs, and the build script refuses to write if it does.

**Table S2.** Every value changed between the upstream analysis output and the submission data set.

| Record | Column | Change | Upstream | Submission | Reason |
|---|---|---|---|---|---|
| C9782 | analytic_class | value recode | drug_or_biological | procedure | Descriptor names a blinded clinical-trial procedure; the sibling code C9783, structurally identical, was already recorded as a procedure. The code stays excluded and no cohort membership or result changes |
| 12 records | reason | value recode | Non-drug wound/skin-substitute item; ret | Wound or skin-substitute item named by t | The upstream text said the code was retained only in sensitivity analysis, which contradicts its membership of the principal 48-code cohort |
| all 288 cells | screen_ disposition | column replaced | core / sensitivity_only | column removed, analytic_cohort_member a | The value sensitivity_only was carried on all 72 cells of the twelve Q-coded items although those items belong to the principal cohort. The ambiguous column is replaced by an unambiguous membership flag |
| S1091 | (row) | value recode | present | removed | The family map listed 49 codes and 42 families because it retained the excluded S-series code; the analytic cohort is 48 codes in 41 families |
| cohort_ definition | definition | value recode | device/device-related non-drug | descriptor-screened supplied-product | Terminology aligned with the manuscript, which assigns no regulatory device status |
| row label | analysis_id | value recode | ..._21_month_window_end | ..._service_period_end | The discarded interval label replaced by the service-period wording used in the manuscript |
| cohort_ definition | definition | value recode | 21-month window end | service-period end | The discarded interval label replaced by the service-period wording used in the manuscript |
| all 49 rows | screen_ disposition | value recode | core / sensitivity_only | column removed, analytic_cohort_member a | Same ambiguity as in the cell matrix: the value sensitivity_only was carried by codes belonging to the principal cohort |

## 3 Data dictionary

Every column of the two central data files is defined below. Both listings run across several pages in a single column, because a table of that length cannot be set inside a column of this width.

| Column | Meaning |
|---|---|
| effective_year | cohort year; the code became effective on 1 April of this year |
| hcpcs_code | five-character HCPCS Level II code |
| long_description_ at_entry | controlling CMS long descriptor in the April file |
| analytic_class | product class assigned by the descriptor screen |
| wound_or_skin_ substitute_item | 1 for the 19 wound or skin-substitute items |
| product_family_ id | identifier shared by codes for one named system or an express parent-and-component relation |
| betos_at_entry | Berenson-Eggers Type of Service category in the April record |
| type_of_service_ at_entry | type-of-service value in the April record |
| pricing_ indicator_at_ entry | raw pricing indicator; 00 not separately priced by Part B, 99 value not established |

continues

Table S3 continued

| Column | Meaning |
|---|---|
| pricing_group | grouping used for the exploratory contrast |
| coverage_code_at_entry | raw coverage indicator; C carrier judgement, I not payable by Medicare, S non-covered by Medicare statute |
| coverage_c | 1 when the coverage indicator is C |
| opps_code_at_entry | Outpatient Prospective Payment System code; empty for every code in this cohort |
| opps_payment_indicator_at_entry | OPPS payment indicator; empty for every code in this cohort |
| code_effective_date | 1 April of the cohort year |
| service_period_end | 31 December of the following year |
| represented_in_either_service_year_file | primary outcome; 1 when a national exact-code row appears in either annual file. Legacy column name from an earlier draft, retained for byte-identity with the frozen output; the endpoint is the one defined in the article and carries no interval claim |
| represented_in_effective_year_file | 1 when the file for service year $Y$ carries a national row |
| represented_in_following_year_file | 1 when the file for service year $Y$+1 carries a national row |
| representation_state | mutually exclusive annual state |
| persistence_across_service_year_files | four-state pattern across the two annual files |
| earlier_represented_service_year | service year of the first file carrying a national row |
| represented_dataset_set | dataset carrying that first row |
| multiple_datasets_in_earlier_represented_year | 1 when more than one dataset carried a row in that year; 0 throughout |
| physician_y_national_aggregate_present | raw field carried through from the source record |
| physician_y_national_row_count | raw field carried through from the source record |
| physician_y_state_row_count | raw field carried through from the source record |
| dme_y_national_aggregate_present | raw field carried through from the source record |
| dme_y_national_row_count | raw field carried through from the source record |
| dme_y_state_row_count | raw field carried through from the source record |
| outpatient_y_national_aggregate_present | raw field carried through from the source record |

continues

Table S3 continued

| Column | Meaning |
|---|---|
| outpatient_y_national_row_count | raw field carried through from the source record |
| outpatient_y_state_row_count | raw field carried through from the source record |
| physician_y_plus_1_national_aggregate_present | raw field carried through from the source record |
| physician_y_plus_1_national_row_count | raw field carried through from the source record |
| physician_y_plus_1_state_row_count | raw field carried through from the source record |
| dme_y_plus_1_national_aggregate_present | raw field carried through from the source record |
| dme_y_plus_1_national_row_count | raw field carried through from the source record |
| dme_y_plus_1_state_row_count | raw field carried through from the source record |
| outpatient_y_plus_1_national_aggregate_present | raw field carried through from the source record |
| outpatient_y_plus_1_national_row_count | raw field carried through from the source record |
| outpatient_y_plus_1_state_row_count | raw field carried through from the source record |
| any_cell_with_multiple_national_partitions | 1 when any cell held more than one legitimate national partition |
| pre_effective_year_exact_code_row_count | negative control; exact-code rows before the effective year |
| pre_effective_year_national_row_present | raw field carried through from the source record |
| usable_throughout_service_period | 1 when the code remained usable to the end of its service period |
| termination_date_in_april_2025_file | termination date recorded in the April 2025 reference file |
| termination_relation_to_service_period | whether termination falls before, on or after the service-period end |
| successor_code | successor recorded in the April 2025 reference file, where one exists |
| endpoint_interpretation | fixed sentence stating what the cell value does and does not mean |

**Table S3.** Columns of Additional file 2, the analytic cohort of 48 code events.

| Column | Meaning |
|---|---|

continues

Table S4 continued

| Column | Meaning |
|---|---|
| effective_year | cohort year of the code |
| hcpcs_code | five-character HCPCS Level II code |
| analytic_class | product class assigned by the descriptor screen |
| pricing_indicator | raw pricing indicator in the April record |
| coverage_code | raw coverage indicator in the April record |
| service_year | service year covered by the queried file |
| service_year_role | 0 for the file covering service year $Y$, 1 for $Y$+1 |
| dataset | queried dataset: physician, dme or outpatient |
| national_exact_code_row_present | 1 when at least one national exact-code row was returned |
| raw_row_count | rows of any geography level returned for the exact code |
| national_row_count | national rows returned; values above one are legitimate partitions |
| state_row_count | state-level rows returned; retained for quality control only |
| unique_state_count | distinct states among those rows |
| national_beneficiary_values | beneficiary counts as returned, never summed across partitions |
| national_service_values | service counts as returned, never summed across partitions |
| national_provider_or_supplier_values | provider or supplier counts as returned, never summed |
| aggregation_caution | flag set when a cell holds more than one national partition |
| interpretation | fixed sentence stating what the cell value does and does not mean |
| analytic_cohort_member | 1 for the 48 code events of the analytic cohort, 0 otherwise; the single authoritative membership flag |

**Table S4.** Columns of Additional file 3, the 288 measurement cells.

Two defects in the upstream analysis output were fixed for the submission data set rather than explained away. Column names such as national_aggregate_within_21_months and first_visibility_channel_set carried an endpoint construct the article no longer uses, and cohort_status used one value, sensitivity_only, both for twelve codes inside the analytic cohort and for one code outside it. The delivered files in supplement/data carry neutral names and an unambiguous disposition; column_mapping.csv lists every rename and corrections.csv every value changed. The upstream output ships unaltered beside them so that the manifest hashes stay verifiable, and the build script asserts that all 4,109 compared values match except the single logged correction.

Two conventions in these files repay attention. A national row count above one is not an error: place of service partitions a national physician aggregate, and supplier rental status partitions a national equipment aggregate, so one code, one service year and one dataset can legitimately carry several national rows. 4 cells did. The binary outcome counts each such cell once and the volume fields are never summed across partitions, because their denominators differ. Second, the columns holding beneficiary, service and provider values are reproduced exactly as returned and are supplied for transparency, not for aggregation.

## 4 Datasets queried

**Table S5.** CMS public-use datasets queried, by claims domain and service year. Every request took the form https://data.cms.gov/data-api/v1/dataset/{id}/data?size=5000&filter[HCPCS_Cd]={code}, with the exact five-character code substituted. Service year $Y-1$ files supplied the negative control.

| Claims domain | Service year | Dataset identifier |
|---|---|---|
| DMEPOS | 2020 | c5da241c-71d2-4b04-9f44-b324ff4b4c01 |
| DMEPOS | 2021 | daedccd6-5881-4997-93cb-ad6f3d1dd2fd |
| DMEPOS | 2022 | a9d1461a-9c63-49e1-8d91-696082e20e47 |
| DMEPOS | 2023 | 87dc62d2-5671-4eb0-9b1a-dfe267a1fa1d |
| DMEPOS | 2024 | 27c150fd-8578-43b1-bba5-6388987e32af |
| Outpatient | 2020 | 31db4293-f75c-4940-bfae-32d554cd164f |
| Outpatient | 2021 | b3537e90-c3c5-4895-914d-71f54123a0a4 |
| Outpatient | 2022 | 6af002ba-b934-4491-96dd-d041989b843c |
| Outpatient | 2023 | 8029af7e-b3d7-486f-ada3-4dbd6a6d8996 |
| Outpatient | 2024 | 04baec39-4a54-400e-824d-8e75251ceda9 |
| Physician | 2020 | 31eb3018-43e0-4259-a0d8-e7a1112ffd08 |
| Physician | 2021 | f00f462f-bb61-48b1-a321-1c50d78ce175 |
| Physician | 2022 | 87304f15-9ed0-41dc-a141-6141a0327453 |
| Physician | 2023 | ddee9e22-7889-4bef-975a-7853e4cd0fbb |
| Physician | 2024 | 6fea9d79-0129-4e4c-b1b8-23cd86a4f435 |

A returned object was accepted only when its HCPCS_Cd field matched the requested code after case and whitespace normalisation. Geography level was read from Rndrng_Prvdr_Geo_Lvl for the physician and outpatient datasets and from Rfrg_Prvdr_Geo_Lvl for the equipment dataset. Only rows at national level satisfied the outcome; state-level rows were retained for quality control. The outpatient rule was never exercised, because no outpatient response returned an exact-code row.

## 5 Sensitivity estimates in full

**Table S6.** Every restriction applied to the cohort, including the three leave-one-cohort- year-out estimates that the article reports only as counts in the running text. Intervals are Wilson 95 per cent intervals computed from the counts in the same row; they describe the precision of a finite cohort and do not treat the cohort as a random sample.

| ID | Analysis | Represented/$N$ | % | 95% Wilson, % | What it changes |
|---|---|---|---|---|---|
| A00 | Analytic cohort of code events | 19/48 | 39.6 | 27.0–53.7 | all 48 descriptor-screened supplied-product code events |
| A01 | Coverage indicator C | 19/46 | 41.3 | 28.3–55.7 | removes A4560 and A7049, which carry indicator S |
| A02 | Decision-document subset | 11/36 | 30.6 | 18.0–46.9 | the 36 codes located in the corresponding second-biannual decision document |
| A03 | Decision-document subset, coverage C | 11/34 | 32.4 | 19.1–49.2 | the same subset without its two indicator-S codes |
| A04 | Excluding wound and skin-substitute items | 8/29 | 27.6 | 14.7–45.7 | removes all 19 A- and Q-prefixed items of that class together |
| A05 | Coverage C and excluding those items | 8/27 | 29.6 | 15.9–48.5 | both restrictions applied together |
| A06 | Codes usable beyond their service-period end | 17/43 | 39.5 | 26.4–54.4 | removes the five K-codes terminating on 31 December 2023 |
| A07 | Excluding the not-otherwise-specified code A4100 | 19/47 | 40.4 | 27.6–54.7 | removes the one code naming a class of products rather than one product |
| A08 | Including S1091, the excluded S-series code | 19/49 | 38.8 | 26.4–52.8 | restores the private-payer series code; its six archived responses hold no exact-code row |
| A09 | Dropping the outpatient dataset | 19/48 | 39.6 | 27.0–53.7 | endpoint computed from the physician and equipment files alone |
| A10 | Product families, any member represented | 17/41 | 41.5 | 27.8–56.6 | a family counts as represented when at least one member is |
| A11 | Product families, all members represented | 15/41 | 36.6 | 23.6–51.9 | a family counts as represented only when every member is |
| A12 | Leaving out the 2021 cohort | 18/40 | 45.0 | 30.7–60.2 | leave-one-cohort-year-out |
| A13 | Leaving out the 2022 cohort | 11/29 | 37.9 | 22.7–56.0 | leave-one-cohort-year-out |
| A14 | Leaving out the 2023 cohort | 9/27 | 33.3 | 18.6–52.2 | leave-one-cohort-year-out |

The spread across these eleven estimates runs from 27.6 to 45.0 per cent. Product composition moves the figure furthest: removing the wound and skin-substitute class lowers it by about twelve percentage points. No single cohort year produces the finding, and the family-level estimate, which counts related codes once, sits close to the headline figure.

## 6 Exploratory contrast by baseline pricing indicator

The article reports this contrast in one short paragraph and states that its instability is the finding. The full set of four analyses is printed here so that a reader can see how far the comparison moves. Risk ratios use a Katz log interval and risk differences a Newcombe unpaired Wilson-score interval, in percentage points. Fisher probabilities are two-sided. None of these quantities is adjusted, and pricing indicator is entangled with code prefix, product family, type of service and Berenson-Eggers group; no multivariable model was fitted, and none could be identified from 48 codes of which 19 were represented.

**Table S7.** Baseline pricing contrast under four cohort definitions. A specific pricing method is compared with indicators 00 or 99 as the reference. The comparison is descriptive: pricing indicator is a record-layout classification measured at code effectiveness, not a fee, a payment or an exposure.

| Cohort definition | Specific | % | 00 or 99 | % | Risk ratio (95% CI) | Risk difference, pp (95% CI) | Fisher $p$ |
|---|---|---|---|---|---|---|---|
| Analytic cohort of code events | 16/28 | 57.1 | 3/20 | 15.0 | 3.81 (1.28–11.35) | 42.1 (14.4–61.2) | 0.0063 |
| Coverage indicator C | 16/28 | 57.1 | 3/18 | 16.7 | 3.43 (1.16–10.11) | 40.5 (11.6–60.1) | 0.0130 |
| Excluding wound and skin-substitute items | 5/9 | 55.6 | 3/20 | 15.0 | 3.70 (1.12–12.24) | 40.6 (4.8–67.9) | 0.0667 |
| Coverage C and excluding those items | 5/9 | 55.6 | 3/18 | 16.7 | 3.33 (1.02–10.92) | 38.9 (2.2–66.7) | 0.0721 |

## 7 Negative control

Each code was queried in the same three datasets for service year $Y - 1$, one year before it could lawfully be used, giving 144 further cells. A positive result would have indicated code reuse, reactivation, or an error in cohort or source assignment. None of the 144 cells returned an exact-code row at any geography level: 0 rows in total. Additional file 4 carries the cell-level record.

Across the 288 cells of the main analysis, the archived responses hold 392 exact-code rows and 29 national code by service year by dataset aggregates.

## 8 Supplementary figure

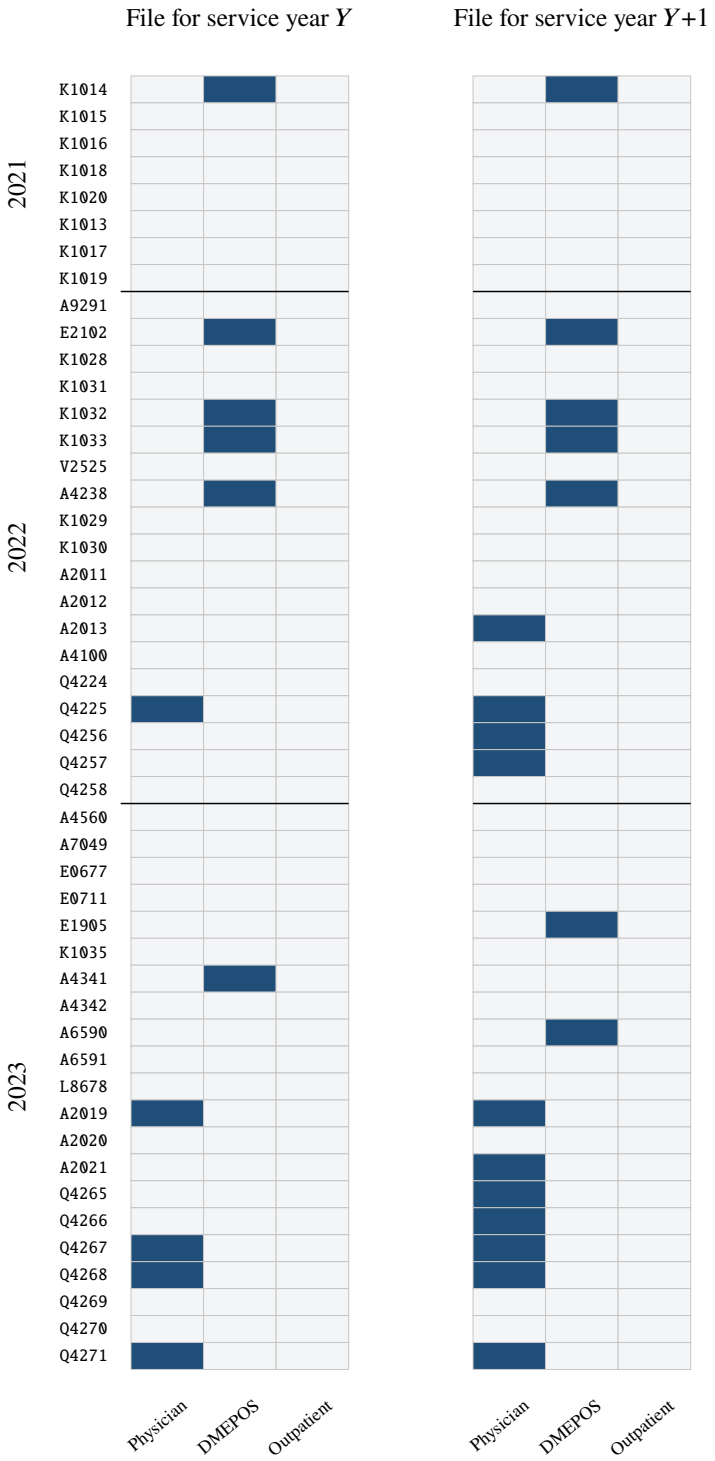



**Figure S1.** All 288 measurement cells for the analytic cohort. A filled cell carries at least one national exact-code row; a pale cell carries none. Codes are ordered within cohort year by the three product groups and then alphanumerically. The left block is the file covering service year $Y$ and the right block the file covering service year $Y + 1$. Cells holding more than one legitimate national partition are filled once. A pale cell is an unresolved state, not an observation of zero use.

## 9 Where the reporting principles are addressed

The table below is *not* a completed STROBE or RECORD checklist, and no such checklist is claimed. It records where the article addresses each of the reporting principles named in its Methods, so that an editor can find them quickly. Items of those statements that presuppose individual-level records do not apply to a design whose unit is an administrative code event.

**Table S8.** Location in the article of each reporting principle named in the Methods.

| Principle | Where it is addressed |
|---|---|
| Source of the data and the population it describes | Methods, source population and cohort construction; Table 2 |
| Rationale and objectives, stated as an answerable question | Background, research question and scope; Table 1 places the question among adjacent designs |
| Complete description of the databases queried | Methods, public data sources and query rule; Table 2; Supplementary Table S4 and Section 4 |
| Codes and algorithms used to define the study population | Methods, source population and cohort construction; Additional files 1 and 2; the full listings, Table S9 |
| Validation of the codes or algorithms | Methods, verification and reproducibility; negative control in Section 7; lifecycle check in Section 10 |
| Definition of the outcome and of every variable | Methods, baseline administrative variables and outcome definitions; Supplementary Table S3 |
| Handling of missing and unresolved values | Methods, query rule; Discussion, why a national row may be missing |
| Quantitative variables and statistical methods | Methods, statistical analysis; Table 5; Supplementary Table in Section 5 |
| Study size and its consequences | Methods, cohort construction; Discussion, strengths and limitations |
| Sensitivity analyses | Results, sensitivity to cohort definition and code dependence; Section 5 |
| Limitations, including data linkage and selection | Discussion, the ceiling on inference and strengths and limitations |
| Access to the data and the analytical code | Declarations, availability of data and materials; Additional files 1 to 8 |

## 10 Full listings

Two code-level listings close the document. Both run in a single column across the remaining pages, because a table of nine and of seven columns cannot be set inside a column of this width. Placing them last keeps the column break from leaving a gap in the body of the document.

Of 103 April additions, 49 passed the descriptor screen and 48 entered the analytic cohort after removal of S1091. 36 of the screened codes were located in the corresponding second-biannual decision document. The 48 codes fall into 41 product families.

**Table S9.** Every code event in the analytic cohort, with its baseline administrative fields and its outcome. Coverage C denotes carrier judgement, I an item not payable by Medicare and S an item non-covered by Medicare statute; none of the three is an adjudicated coverage decision. Pricing indicator 00 means not separately priced by Part B and 99 means value not established.

| Year | Code | Product class | Cov. | Pr. | Family | State | First | Persist. |
|---|---|---|---|---|---|---|---|---|
| 2021 | K1013 | component or supply | C | 99 | 2021 K1013 | neither file | – | neither |
| 2021 | K1014 | prosthetic or orthotic | C | 99 | 2021 K1014 | file for *Y* | DMEPOS | both |
| 2021 | K1015 | prosthetic or orthotic | C | 99 | 2021 K1015 | neither file | – | neither |
| 2021 | K1016 | stand-alone device | C | 99 | 2021 monarch etns | neither file | – | neither |
| 2021 | K1017 | component or supply | C | 99 | 2021 monarch etns | neither file | – | neither |
| 2021 | K1018 | stand-alone device | C | 99 | 2021 cala tremor | neither file | – | neither |
| 2021 | K1019 | component or supply | C | 99 | 2021 cala tremor | neither file | – | neither |
| 2021 | K1020 | stand-alone device | C | 99 | 2021 K1020 | neither file | – | neither |
| 2022 | A2011 | wound product, A prefix | C | 13 | 2022 A2011 | neither file | – | neither |
| 2022 | A2012 | wound product, A prefix | C | 13 | 2022 A2012 | neither file | – | neither |
| 2022 | A2013 | wound product, A prefix | C | 13 | 2022 A2013 | file for *Y*+1 only | Physician | *Y*+1 only |
| 2022 | A4100 | wound product, A prefix | C | 13 | 2022 A4100 | neither file | – | neither |
| 2022 | A4238 | component or supply | C | 34 | 2022 adjunctive cgm | file for *Y* | DMEPOS | both |
| 2022 | A9291 | software device | C | 99 | 2022 A9291 | neither file | – | neither |
| 2022 | E2102 | stand-alone device | C | 32 | 2022 adjunctive cgm | file for *Y* | DMEPOS | both |
| 2022 | K1028 | stand-alone device | C | 99 | 2022 exciteosa | neither file | – | neither |
| 2022 | K1029 | component or supply | C | 99 | 2022 exciteosa | neither file | – | neither |
| 2022 | K1030 | component or supply | C | 99 | 2022 K1030 | neither file | – | neither |
| 2022 | K1031 | stand-alone device | C | 99 | 2022 koya dayspring | neither file | – | neither |
| 2022 | K1032 | stand-alone device | C | 99 | 2022 koya dayspring | file for *Y* | DMEPOS | both |
| 2022 | K1033 | stand-alone device | C | 99 | 2022 koya dayspring | file for *Y* | DMEPOS | both |
| 2022 | Q4224 | skin substitute, Q prefix | C | 51 | 2022 Q4224 | neither file | – | neither |
| 2022 | Q4225 | skin substitute, Q prefix | C | 51 | 2022 Q4225 | file for *Y* | Physician | both |
| 2022 | Q4256 | skin substitute, Q prefix | C | 51 | 2022 Q4256 | file for *Y*+1 only | Physician | *Y*+1 only |
| 2022 | Q4257 | skin substitute, Q prefix | C | 51 | 2022 Q4257 | file for *Y*+1 only | Physician | *Y*+1 only |
| 2022 | Q4258 | skin substitute, Q prefix | C | 51 | 2022 Q4258 | neither file | – | neither |
| 2022 | V2525 | stand-alone device | C | 99 | 2022 V2525 | neither file | – | neither |
| 2023 | A2019 | wound product, A prefix | C | 13 | 2023 A2019 | file for *Y* | Physician | both |
| 2023 | A2020 | wound product, A prefix | C | 13 | 2023 A2020 | neither file | – | neither |

Table S9 continued

| Year | Code | Product class | Cov. | Pr. | Family | State | First | Persist. |
|---|---|---|---|---|---|---|---|---|
| 2023 | A2021 | wound product, A prefix | C | 13 | 2023 A2021 | file for $Y$+1 only | Physician | $Y$+1 only |
| 2023 | A4341 | component or supply | C | 37 | 2023 intraurethral drainag | file for $Y$ | DMEPOS | $Y$ only |
| 2023 | A4342 | component or supply | C | 37 | 2023 intraurethral drainag | neither file | – | neither |
| 2023 | A4560 | stand-alone device | S | 00 | 2023 A4560 | neither file | – | neither |
| 2023 | A6590 | component or supply | C | 34 | 2023 A6590 | file for $Y$+1 only | DMEPOS | $Y$+1 only |
| 2023 | A6591 | component or supply | C | 34 | 2023 A6591 | neither file | – | neither |
| 2023 | A7049 | stand-alone device | S | 00 | 2023 A7049 | neither file | – | neither |
| 2023 | E0677 | stand-alone device | C | 36 | 2023 E0677 | neither file | – | neither |
| 2023 | E0711 | stand-alone device | C | 00 | 2023 E0711 | neither file | – | neither |
| 2023 | E1905 | software device | C | 46 | 2023 E1905 | file for $Y$+1 only | DMEPOS | $Y$+1 only |
| 2023 | K1035 | diagnostic device | C | 00 | 2023 K1035 | neither file | – | neither |
| 2023 | L8678 | component or supply | C | 38 | 2023 L8678 | neither file | – | neither |
| 2023 | Q4265 | skin substitute, Q prefix | C | 51 | 2023 Q4265 | file for $Y$+1 only | Physician | $Y$+1 only |
| 2023 | Q4266 | skin substitute, Q prefix | C | 51 | 2023 Q4266 | file for $Y$+1 only | Physician | $Y$+1 only |
| 2023 | Q4267 | skin substitute, Q prefix | C | 51 | 2023 Q4267 | file for $Y$ | Physician | both |
| 2023 | Q4268 | skin substitute, Q prefix | C | 51 | 2023 Q4268 | file for $Y$ | Physician | both |
| 2023 | Q4269 | skin substitute, Q prefix | C | 51 | 2023 Q4269 | neither file | – | neither |
| 2023 | Q4270 | skin substitute, Q prefix | C | 51 | 2023 Q4270 | neither file | – | neither |
| 2023 | Q4271 | skin substitute, Q prefix | C | 51 | 2023 Q4271 | file for $Y$ | Physician | both |

## Column mapping

**Table S10.** Every column renamed between the upstream analysis output and the submission data set. Values were carried across unchanged.

| Upstream file | Upstream column | Submission column |
|---|---|---|
| screening_ledger.csv | cohort_status | screen_disposition |
| principal_code_ledger.csv | follow_up_window_end | service_period_end |
| principal_code_ledger.csv | national_aggregate_within_21_months | represented_in_either_service_year_file |
| principal_code_ledger.csv | national_aggregate_in_effective_year | represented_in_effective_year_file |
| principal_code_ledger.csv | national_aggregate_in_following_year | represented_in_following_year_file |
| principal_code_ledger.csv | three_state_outcome | representation_state |
| principal_code_ledger.csv | persistence_category | persistence_across_service_year_files |
| principal_code_ledger.csv | first_visibility_calendar_year | earlier_represented_service_year |
| principal_code_ledger.csv | first_visibility_channel_set | represented_dataset_set |
| principal_code_ledger.csv | multiple_channels_at_first_visibility | multiple_datasets_in_earlier_represented_year |
| principal_code_ledger.csv | any_multiple_national_partition_cell | any_cell_with_multiple_national_partitions |
| principal_code_ledger.csv | y_minus_1_exact_code_row_count_all_channels | pre_effective_year_exact_code_row_count |
| principal_code_ledger.csv | y_minus_1_national_aggregate_present_any_channel | pre_effective_year_national_row_present |
| principal_code_ledger.csv | active_for_entire_21_month_window | usable_throughout_service_period |
| principal_code_ledger.csv | termination_relation_to_window | termination_relation_to_service_period |
| principal_code_year_channel_matrix.csv | cohort_status | screen_disposition |
| principal_code_year_channel_matrix.csv | calendar_year | service_year |
| principal_code_year_channel_matrix.csv | follow_up_year | service_year_role |
| principal_code_year_channel_matrix.csv | channel | dataset |
| principal_code_year_channel_matrix.csv | national_exact_code_aggregate_present | national_exact_code_row_present |
| y_minus_1_check.csv | negative_control_year | pre_effective_service_year |
| y_minus_1_check.csv | channel | dataset |
| y_minus_1_check.csv | national_exact_code_aggregate_present | national_exact_code_row_present |
| code_lifecycle_check.csv | follow_up_window_end | service_period_end |
| code_lifecycle_check.csv | active_for_entire_21_month_window | usable_throughout_service_period |
| code_lifecycle_check.csv | termination_relation_to_window | termination_relation_to_service_period |

| Upstream file | Upstream column | Submission column |
|---|---|---|
| final_analysis_summary.csv | observed_by_21_months | represented |
| final_analysis_summary.csv | observed_by_21_months_percent | represented_percent |
| final_analysis_summary.csv | visible_in_effective_year_including_both | represented_in_effective_year_file |
| final_analysis_summary.csv | first_visible_in_following_year | represented_in_following_year_file_only_count |
| final_analysis_summary.csv | no_national_aggregate_in_either_year | not_represented_in_either_file |
| final_analysis_summary.csv | visible_in_both_years | represented_in_both_files |
| final_analysis_summary.csv | visible_in_effective_year_only | represented_in_effective_year_file_only |
| final_analysis_summary.csv | visible_in_following_year_only | represented_in_following_year_file_only |
| final_analysis_summary.csv | first_visibility_physician | represented_dataset_physician |
| final_analysis_summary.csv | first_visibility_dme | represented_dataset_dmepos |
| final_analysis_summary.csv | first_visibility_outpatient | represented_dataset_outpatient |
| final_analysis_summary.csv | multiple_channels_at_first_visibility | multiple_datasets_in_earlier_represented_year |
| final_analysis_summary.csv | cohort_2021_visible_in_y | cohort_2021_represented_in_effective_year_file |
| final_analysis_summary.csv | cohort_2021_first_visible_in_y_plus_1 | cohort_2021_represented_in_following_year_file_only |
| final_analysis_summary.csv | cohort_2021_both_years | cohort_2021_both_files |
| final_analysis_summary.csv | cohort_2021_y_only | cohort_2021_effective_year_file_only |
| final_analysis_summary.csv | cohort_2021_y_plus_1_only | cohort_2021_following_year_file_only |
| final_analysis_summary.csv | cohort_2022_visible_in_y | cohort_2022_represented_in_effective_year_file |
| final_analysis_summary.csv | cohort_2022_first_visible_in_y_plus_1 | cohort_2022_represented_in_following_year_file_only |
| final_analysis_summary.csv | cohort_2022_both_years | cohort_2022_both_files |
| final_analysis_summary.csv | cohort_2022_y_only | cohort_2022_effective_year_file_only |
| final_analysis_summary.csv | cohort_2022_y_plus_1_only | cohort_2022_following_year_file_only |
| final_analysis_summary.csv | cohort_2023_visible_in_y | cohort_2023_represented_in_effective_year_file |
| final_analysis_summary.csv | cohort_2023_first_visible_in_y_plus_1 | cohort_2023_represented_in_following_year_file_only |
| final_analysis_summary.csv | cohort_2023_both_years | cohort_2023_both_files |
| final_analysis_summary.csv | cohort_2023_y_only | cohort_2023_effective_year_file_only |
| final_analysis_summary.csv | cohort_2023_y_plus_1_only | cohort_2023_following_year_file_only |
| y_minus_1_source_manifest.csv | negative_control_year | pre_effective_service_year |
| y_minus_1_source_manifest.csv | channel | dataset |
| screening_ledger.csv | cohort_status | screen_disposition |
| screening_ledger.csv | (derived from principal_code_ledger.csv) | analytic_cohort_member |
| screening_ledger.csv | cohort_status | decision_document_subset |
| screening_ledger.csv | cohort_status | provenance_upstream_cohort_status |

## Code lifecycle

Termination and action fields were read from the April 2025 full HCPCS reference file. Under the CMS record layout the termination date is the last date on which a code may be used by Medicare providers, so a code terminating on 31 December of year $Y + 1$ remained usable throughout its service period. All 48 codes did.

**Table S11.** Lifecycle adjudication for every code event. The final column records whether the code remained usable to the end of its service period.

| Year | Code | Period end | Termination | Relation to period | Successor | Active |
|---|---|---|---|---|---|---|
| 2021 | K1013 | 20221231 | 20231231 | termination after window | A4457 | 1 |
| 2021 | K1014 | 20221231 | 20231231 | termination after window | L5615 | 1 |
| 2021 | K1015 | 20221231 | 20231231 | termination after window | L3161 | 1 |
| 2021 | K1016 | 20221231 | 20231231 | termination after window | E0733 | 1 |
| 2021 | K1017 | 20221231 | 20231231 | termination after window | A4541 | 1 |
| 2021 | K1018 | 20221231 | 20231231 | termination after window | E0734 | 1 |
| 2021 | K1019 | 20221231 | 20231231 | termination after window | A4542 | 1 |
| 2021 | K1020 | 20221231 | 20231231 | termination after window | E0735 | 1 |
| 2022 | A2011 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | A2012 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |

Table S11 continued

| Year | Code | Period end | Termination | Relation to period | Successor | Active |
|---|---|---|---|---|---|---|
| 2022 | A2013 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | A4100 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | A4238 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | A9291 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | E2102 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | K1028 | 20231231 | 20231231 | termination on window end last usable day | E0492 | 1 |
| 2022 | K1029 | 20231231 | 20231231 | termination on window end last usable day | E0493 | 1 |
| 2022 | K1030 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | K1031 | 20231231 | 20231231 | termination on window end last usable day | E0681 | 1 |
| 2022 | K1032 | 20231231 | 20231231 | termination on window end last usable day | E0678 | 1 |
| 2022 | K1033 | 20231231 | 20231231 | termination on window end last usable day | E0679 | 1 |
| 2022 | Q4224 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | Q4225 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | Q4256 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | Q4257 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | Q4258 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2022 | V2525 | 20231231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | A2019 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | A2020 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | A2021 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | A4341 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | A4342 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | A4560 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | A6590 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | A6591 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | A7049 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | E0677 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | E0711 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | E1905 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | K1035 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | L8678 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | Q4265 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | Q4266 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | Q4267 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | Q4268 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | Q4269 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | Q4270 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |
| 2023 | Q4271 | 20241231 | none recorded | no termination date in april 2025 file | – | 1 |