National Dementia Workforce Study: Community Clinician Wave 1 Sample Frame, At-A-Glance

The National Dementia Workforce Study aims to survey a nationally representative sample of clinicians who care for people living with dementia. The clinicians sought for Wave 1 of the survey included primary care physicians, neurologists, psychiatrists, primary care nurse practitioners, mental health nurse practitioners, and non-surgical physician assistants who provide care to Medicare beneficiaries with dementia in outpatient or residential settings.

This document describes the data and methods used to construct the sample frame, and the SAS code used to construct the frame are available in the accompanying file (NDWS_CC_Sample_Frame_W1_20250611.zip)[1].

Data

To develop a nationwide sample frame of clinicians who care for people living with dementia in fee-for-service Medicare and Medicare Advantage, we required multiple data sources and needed to account for differences in data availability. For this reason, the Wave 1 sample frame used data from calendar year 2021 (Medicare fee-for-service claims, Medicare Advantage encounter data, Part D drug event file, and Minimum Data Set assessments) and federal fiscal year 2023 (fee-for-service claims and Part D drug event file). The sample frame also used other publicly available and proprietary data sets (see Table 1).

Methods

The overall approach to identifying the sample frame is available in the overview of methods to develop sample frame section. The process involves five broad steps, and within each of the broad steps, there are often multiple sub-steps (and multiple SAS programs) to implement the required logic to identify the sample frame. The following are the five broad steps, and each has a link to the relevant section that provides additional details:

1. Identify the primary taxonomy of all individual clinicians nationwide and their most recent practice address.
2. Identify Medicare-enrolled beneficiaries with dementia and their enrollment characteristics, specifically whether they were dually eligible for Medicare and Medicaid or eligible for a Part D low-income drug subsidy.
3. Identify clinicians who served the beneficiaries with dementia and the settings in which they practiced.
4. Identify the unique set of clinicians with relevant primary taxonomy and licensure and create the National Provider Identifier–level claims-based variables needed for sampling, such as number of Medicare beneficiaries with dementia served and in which settings.
5. Geocode clinicians' practice addresses to determine whether they worked in rural regions and apply a final set of sample restrictions (for example, dropping clinicians who worked in U.S. territories or elsewhere abroad).

[1] For readers without SAS, the programs can be read in any text editor. ***Mac users can open the files with TextEdit but may need to go to the "Privacy & Security" menu in their System Settings to do so***.

A. Background

The National Dementia Workforce Study (NDWS) aims to survey a nationally representative sample of clinicians caring for persons living with dementia (PLWD). Clinicians sought for Wave 1 of the survey included primary care physicians, neurologists, psychiatrists, primary care nurse practitioners, mental health nurse practitioners, and non-surgical physician assistants who provide care in outpatient or residential settings. To develop the Wave 1 sample frame, we used Medicare fee-for-service (FFS) claims, Medicare Advantage (MA) encounter data, Medicare Part D event files,[1] and Minimum Data Set (MDS) data, along with publicly available data, including data for primary taxonomy from the National Plan and Provider Enumeration System (NPPES). To geocode clinicians' practice locations, we obtained proprietary address data for the clinicians from a commercial vendor, IQVIA OneKey, and mapped these addresses to Rural-Urban Commuting Area (RUCA) codes using publicly available data from the U.S. Census Bureau and the U.S. Department of Agriculture.

In this paper and accompanying documentation (SAS code used to develop the sample frame), we provide an overview of the data and methods used to develop the sample frame (Section B) and descriptions of the code used to develop the sample frame (Section C).

B. Data and methods to develop the community clinician sample frame

B.1 Data

Table 1 describes the data sources used to construct the community clinician sample frame for Wave 1. For the Medicare FFS claims, MA encounter, MDS, and Part D event data, we generally used the most recent files available at the time the work was conducted. However, one challenge was that the most recent period of data available for each data source was not the same. Specifically, the most recent FFS data covered federal fiscal year 2023 (FY 2023; October 1, 2022, to September 30, 2023), whereas the most recent Part D data covered calendar year (CY) 2022 and the most recent MA encounter and MDS data covered CY 2021.

Table 1. Data sources to construct the community clinician sample frame for Wave 1 of the NDWS

Data (years)	Used in sample frame construction to:	Access to data
Medicare service use data: FFS claims (CY 2021 and FY 2023), MA encounter data (CY 2021), and Part D event files (CY 2021 and 4th quarter of CY 2022[a])	• Identify Medicare beneficiaries with dementia • Determine the settings where clinicians provided care to beneficiaries with dementia • Count the number of beneficiaries living with dementia that each clinician saw	Accessed via the Chronic Conditions Warehouse (CCW) Virtual Research Data Center (VRDC) under approved data use agreement (DUA)

[1] The Part D event files contain information on prescription fills made by Medicare beneficiaries with Part D coverage enrolled in FFS or MA.

Minimum Data Set (CY 2021)	• Identify Medicare-enrolled nursing home residents with dementia not identified in claims or encounter data	Accessed via the CCW VRDC under approved DUA
Medicare enrollment data (CY 2021 and FY 2023)	• Count the number of Medicare beneficiaries with dementia in each clinician's panel who were dually eligible for Medicare and Medicaid or eligible for a low-income drug subsidy	Accessed via the CCW VRDC under approved DUA
National Plan and Provider Enumeration System (NPPES) (CY 2021, FY 2023, and April 2024)	• Identify clinicians' primary taxonomy and their practice address[b]	Publicly available from the Centers for Medicare & Medicaid Services (CMS) website: https://www.cms.gov/medicare/regulations-guidance/administrative-simplification/data-dissemination
Provider Enrollment, Chain, and Ownership System (PECOS) (2023)	• Help identify clinicians who practice in U.S. territories or elsewhere outside of the U.S.	Publicly available from the CMS website: https://data.cms.gov/provider-characteristics/medicare-provider-supplier-enrollment/medicare-fee-for-service-public-provider-enrollment
Zip code tabulation area (ZCTA) relationship files (2010)	• Map each National Provider Identifier's most frequent practice zip code from the claims data to its associated county	Publicly available from the U.S. Census Bureau: https://www.census.gov/geographies/reference-files/time-series/geo/relationship-files.2010.html#par_textimage_674173622
Medicare Geographic Variations Database (2021 and 2022)	• Calculate the percentage of Medicare beneficiaries enrolled in FFS in each county nationwide	Publicly available from the CMS website: https://data.cms.gov/summary-statistics-on-use-and-payments/medicare-geographic-comparisons/medicare-geographic-variation-by-national-state-county
IQVIA OneKey data (2024)	• Geocode clinicians' addresses	Purchased from IQVIA
U.S. Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER)/line shapefile (2010)	• Map clinicians' practice address to census tract polygon coordinates	Publicly available from the U.S. Census Bureau: https://www.census.gov/geographies/mapping-files.html
Rural–Urban Commuting Area Codes (RUCA) file (2010)	• Assign clinicians' practice address to a RUCA code (based on geocoded census tract)	Publicly available from the U.S. Department of Agriculture website: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/

[a] The fourth quarter of CY 2022 is the first quarter of FY 2023, meaning that we had Part D data for one quarter of FY 2023.

[b] Our primary data source for practice address was the IQVIA OneKey data. However, for any clinicians without addresses in the IQVIA data, we used address data from NPPES (and PECOS) to determine whether the clinician worked in a U.S. territory or elsewhere abroad, and if not, backfilled their address information using their address from NPPES.

B.2 Overview of methods to develop sample frame

To build the sample frame, we first identified the universe of Medicare beneficiaries with dementia who received care in CY 2021 (FFS and MA) and FY 2023 (FFS only) using the Bynum standard algorithm,[2] updated to capture the additional specifiers available in FY 2023 (Appendix Table 1; for example, F01.511 [Vascular dementia, unspecified severity, with agitation]). We did not include diagnoses included on MA chart review records submitted by managed care organizations because equivalent data are not available for FFS claims data. We also identified nursing home residents with dementia based on MDS assessments (items I4200 and I4800) from CY 2021.

Having comprehensively identified Medicare beneficiaries with dementia using all FFS claims, MA encounter records, and MDS assessment data available, our next step was identifying the clinicians who provided care to these individuals and the settings in which these clinicians delivered care to beneficiaries with dementia. We identified the individual clinicians by searching the carrier files (that is, where clinician services are billed to Medicare) and the outpatient files, the latter limited to claims from Federally Qualified Health Centers (FQHCs), rural health centers (RHCs), and type II critical access hospitals (CAHs)[3]—that is, claims and encounters for outpatient services where there is no corresponding carrier record because the clinic claim covers both the facility and professional components.

Because nurse practitioners and physician assistants might bill for their services under a supervising physician and, therefore, their role is likely underrepresented in claims data,[4] we used the Medicare Part D event files (CY 2021 and the last quarter of CY 2022 [Q1 FY 2023]) to identify clinicians who prescribed to beneficiaries with dementia but were not otherwise found in the carrier and outpatient claims or encounter data in CY 2021 or FY 2023.

Once we identified the clinicians providing care, we used the place of service variable (carrier claims) and revenue center codes (outpatient claims) to assign each visit to the following

[2] Grodstein F., C.-H. Chang, A.W. Capuano, M.C. Power, D.X. Marquez, L.L. Barnes, et al. "Identification of Dementia in Recent Medicare Claims Data, Compared With Rigorous Clinical Assessments." *The Journals of Gerontology, Series A,* vol. 77, no. 6, 2021, pp. 1272–1278. https://doi.org/10.1093/gerona/glab377.

[3] Type II CAHs bill for both the facility and professional components of a visit. In contrast, type I CAHs bill only the facility component of a visit. For more information, see Centers for Medicare & Medicaid Services. "Information for Critical Access Hospitals." Medical Learning Network, 2024. https://www.cms.gov/files/document/mln006400-information-critical-access-hospitals.pdf.

[4] For example, see Patel, S.Y., H.A. Huskamp, A.B. Frakt, D.I. Auerbach, H.T. Neprash, M.L. Barnett, et al. "Frequency of Indirect Billing to Medicare for Nurse Practitioner and Physician Assistant Office Visits." *Health Affairs,* vol. 41, no. 6, 2022. https://doi.org/10.1377/hlthaff.2021.01968.

settings: outpatient, inpatient, residential, Part D prescriber, or all other settings (Appendix Table 2). For clinicians who had claims for beneficiaries with dementia in the Part D event data only, we also pulled their carrier and outpatient claims for beneficiaries without dementia (if any) to reassign their setting.[5] That is, if they submitted no carrier or outpatient claims for beneficiaries with dementia, we used the settings where they served beneficiaries without dementia, assuming the clinicians would serve beneficiaries with dementia in the same setting. The setting for clinicians found in Part D claims for Medicare beneficiaries with dementia but not otherwise found in carrier or outpatient claims remained "Part D prescriber."

Having identified clinicians serving beneficiaries with dementia, we then limited the sample frame to the specific clinician specialties and licensures of interest—primary care physicians (including geriatrics and hospice/palliative care), neurologists, psychiatrists, primary care nurse practitioners (including geriatrics), mental health nurse practitioners, and non-surgical physician assistants. We did so using their primary taxonomy as classified in the NPPES (Appendix Table 3). We also limited the file to clinicians whose National Provider Identifier (NPI) was still active according to the April 2024 NPPES. With these criteria, we identified 580,134 clinicians in CY 2021 and 537,232 clinicians in FY 2023. To this set of clinicians, we added 308 clinicians from CY 2021 with relevant a taxonomy who (1) were still active in April 2024, (2) had cared for beneficiaries with dementia identified in the MDS only, and (3) had not been identified in the claims or encounter data as serving any other beneficiaries with dementia.

After examining the overlap of clinicians found during the two 12-month periods, we dropped 553,135 clinicians from the CY 2021 file for two reasons: (1) they were also found in the FY 2023 file or (2) they served FFS beneficiaries with dementia in CY 2021 only and were not found in the FY 2023 file and thus were likely retired. After this step, the only clinicians retained in the sample frame from CY 2021 were those who served MA-enrolled beneficiaries with dementia or those who treated beneficiaries with dementia only observed in the MDS. We retained these clinicians in the frame given they were likely still clinically active, but 2023 MA and MDS data were not available to confirm this at the time of sample frame construction.

We then constructed several measures from the claims and encounter data to help guide sampling. These included: (1) the number of beneficiaries with dementia cared for by the clinician, (2) the percentage of the clinician panel comprising beneficiaries with dementia who were ever dually eligible for Medicare and Medicaid or eligible for a Part D low-income drug subsidy (DE/LIS) during the year, and (3) the settings in which the clinician provided care to beneficiaries with dementia. We then purchased information on primary practice address for this sample from a commercial vendor (IQVIA OneKey) and geocoded the clinician's practice address

[5] We identified claims from these clinicians based on the performing NPI field (FFS carrier line file), servicing NPI field (MA carrier header file), and attending NPI field (outpatient header file). These fields were nearly universally populated in the FFS data. In the MA encounter data, which had higher rates of missing data in these fields, we backfilled missing servicing NPI with the organizational NPI in the carrier file. Similarly, in the outpatient data, we backfilled missing attending NPI with rendering NPI, if not missing, else with organizational NPI.

to determine census tract and its associated RUCA codes.[6] The geocoded file was used to construct another variable used for sampling, specifically (4) whether the clinician worked in a rural region. Finally, we excluded clinicians who had encounters only in inpatient settings (n = 37,867), worked only in "all other settings" (for example, emergency rooms or community mental health centers; n = 31,649), or worked outside of the United States (n = 2,840). The final sample frame included 492,183 clinicians (Appendix Table 4).

C. Code used to construct the community clinician sample frame

The process for identifying the sample frame includes the following broad steps:

1. Use NPPES data to identify the primary taxonomy of all individual clinicians nationwide and their most recent practice address.
2. Identify Medicare-enrolled beneficiaries with dementia and determine whether they are DE/LIS.
3. Identify clinicians who served the beneficiaries with dementia and the settings in which they practice.
4. Identify the unique set of clinicians across CY 2021 and FY 2023 with relevant licensure and primary taxonomy and create variables needed for sampling.
5. Finalize the sample frame after geocoding clinician's practice addresses to determine if they worked in rural regions and applying a final set of sample restrictions (for example, dropping clinicians who worked in U.S. territories or elsewhere abroad).

Within each step, there are often multiple programs and steps that implement the required logic to identify the final sample frame. This is particularly true because the sample frame uses data from CY 2021 (FFS claims, MA encounter data, Part D drug event file, and MDS assessments) and FY 2023 (FFS claims and Part D drug event file). The steps are described below. The SAS code is available in the accompanying zip file, NDWS_CC_Sample_Frame_W1_20250611.zip. For readers without SAS, the programs can be read in any text editor. ***Mac users can open the files with TextEdit but may need to go to the "Privacy & Security" menu in their System Settings to do so***.

1. Use NPPES data to identify the primary taxonomy of all individual clinicians nationwide and their most recent practice address

- 010_import_taxonomy_xwalk.sas. This program reads an Excel worksheet provided by the University of Michigan into SAS with a list of taxonomy codes and indicators for licensure type (MD/DO, nurse practitioner [NP], or physician assistant [PA]) and whether the taxonomy is of interest for the community clinician survey (see Appendix Table 3 for a list

[6] For more information on RUCA codes, see https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/documentation.

of taxonomy codes and licensures of interest). The SAS data set of taxonomy codes is used in the next step to identify clinicians with qualifying specialties.

- 020_npi.sas. This program reads in NPPES data from April 2024, FY 2023, and CY 2021, and does the following for each year: 1) limits to records for individual clinicians (that is, excludes organizational NPIs) and relevant variables; 2) finds the primary taxonomy code on each record, where available; 3) finds the most recent record per NPI, separately for CY 2021 and FY 2023, and flags records with non-missing primary taxonomy code in April 2024; 4) merges the three files (April 2024, FY 2023, and CY 2021) by NPI so that we have one record for each individual NPI who was active in any of these periods; 4) creates flags for whether the NPI was active in each period and whether their primary taxonomy was one of interest for the survey based on the file from the 010 program; and 5) creates a flag to identify the "priority practice address" for each NPI (for this variable, we prioritized practice address over business mailing address and data from the most recent year versus prior years)

2. Identify Medicare-enrolled beneficiaries with dementia and determine whether they are DE/LIS

- 050_create_dx_codeset.sas. This program creates a SAS data set containing all diagnosis codes used to identify Medicare beneficiaries with dementia.

- 100_create_ffs_finders.sas. This program applies the Bynum-Standard algorithm[2] (any inpatient, skilled nursing facility [SNF], home health agency, or hospice claim or two or more carrier or FQHC/RHC/type II CAH claims at least seven days apart with a relevant dementia diagnosis code in any position) to Medicare FFS claims to identify Medicare FFS beneficiaries with dementia (see Appendix Table 1 for a list of diagnosis codes and rules). We run this for both CY 2021 and FY 2023. As part of the outpatient file processing to identify claims from FQHCs, RHCs, and type II CAHs, the code also creates flags for the settings in which the FQHC, RHC, or type II CAH visit occurred (clinic, SNF, or home or other settings) based on revenue center codes (this will be used downstream in the 210 program described below). The code outputs all FQHC, RHC, and type II CAH claims for each year (CY 2021 and FY 2023) into separate output files for use in the 210 program. This program outputs beneficiary-level files for the two time periods with all beneficiaries who met the Bynum-Standard dementia definition based on FFS claims.

- 101_flag_dementia.sas. This program reads in managed care encounter records for CY 2021 and flags those with a dementia diagnosis code using the set of diagnosis codes read into the file created in the 050 program described above. Similar to the 100 program described above, the code processes outpatient encounter records to identify claims from FQHCs, RHCs, and type II CAHs; creates flags for the settings in which the FQHC, RHC, or type II CAH visit occurred; and outputs all FQHC, RHC, and type II CAH claims into an output file for use in the 102 and 310 programs described below.

- 102_bene_lvl_dementia.sas. This program reads in the encounter data output by the 101 program by encounter record type and retains final action encounter records with a dementia diagnosis (as described above, we exclude chart review records from the algorithm to ensure that we apply the Bynum-Standard algorithm similarly on FFS claims and managed care encounter data). It outputs a beneficiary-level file containing all beneficiaries who met the Bynum-Standard dementia definition based on their CY 2021 encounter data.

- 103_dementia_benes.sas. This program flags resident assessments from the CY 2021 MDS where the resident had an active diagnosis of Alzheimer's disease or other dementia (i4200_alzhmr_cd = 1 or i4800_dmnt_cd = 1) and outputs a resident-level file with an indicator variable for whether they have dementia.

- 105_ffs_ma_mds_dementia_benes.sas. In this program, the code merges the three beneficiary-level output files from the 100 (FFS), 102 (MA), and 103 (MDS) programs by unique beneficiary identifier. The code drops beneficiaries in the MDS without dementia—that is, who did not merge to a record for a beneficiary with dementia based on FFS or MA encounter data—and retains all MDS residents with dementia.

- 110_add_mbsf_info_baseyr.sas. This program pulls Master Beneficiary Summary File (MBSF) data for beneficiaries with dementia based on the CY 2021 data (in the output file from the 105 program), specifically Medicare enrollment and demographic data. The code creates several indicator variables; the relevant indicator variable for the sample frame is whether the beneficiary was DE/LIS in any month during the year.[7] The code drops beneficiaries whose death date indicated they died before the start of CY 2021 and beneficiaries who were not found in the MBSF.

- 120_add_mbsf_info_recentyr.sas. This program pulls MBSF data for beneficiaries with dementia based on the FY 2023 data (in the output file from the 100 program). It generally mirrors the 110 program except that it pulls monthly eligibility data from two distinct, yearly MBSFs. Specifically, from the CY 2022 MBSF, the program pulls monthly data on DE/LIS status from October through December. Similarly, it pulls monthly eligibility data for January–September 2023 from the CY 2023 MSBF.

3. Identify clinicians who served the beneficiaries with dementia and the settings in which they practice

[7] The program also constructed separate indicator variables for dual eligibility status and low-income drug subsidy status as well as various indicators for race and ethnicity, but we ultimately did not use these variables to develop the sample frame.

- 210 find npis.sas. This program has four key steps and involves identifying and processing the carrier and FQHC/RHC/type II CAH FFS claims for FFS beneficiaries with dementia to identify the NPIs who served them. Ultimately, the program outputs an NPI-level file with information about the setting(s) in which each NPI served their FFS dementia patients, the number and types of FFS beneficiaries with dementia served, and information about practice zip code. The program is run on both CY 2021 and FY 2023 data. The steps are as follows:
    - Step 1: Process carrier data. This step identifies FFS carrier header and line-level data for beneficiaries with dementia, retaining only a small set of variables needed for the processing (for example, diagnosis codes from the header and performing clinician NPI, place of service, and the NPI's practice zip code from the line file). It joins the headers and lines and retains only the first line per claim, so that each claim is only counted once. The code also merges eligibility and demographic characteristics from the 120 file onto the claims. The code uses the diagnosis codes to create an indicator variable for whether the claim included a dementia diagnosis, per the Bynum-Standard code set. It also creates indicator variables from the place of service variable to describe the setting in which the service was delivered—either outpatient, inpatient, residential, or other (see Appendix Table 2). The code also creates a setting variable for "Part D prescriber," which is set to 0 for all records.
    - Step 2: Process Part D drug event data. This step identifies Part D claims for FFS beneficiaries with dementia. It merges eligibility and demographic characteristics from the 120 file onto the Part D claims. The code also creates the same set of indicator variables for settings as those in Step 1, including a setting for "Part D prescriber." In this file, all claims have the "Part D prescriber" flag set to 1 and all other setting variables set to 0.
    - Step 3: Process FFS FQHC, RCH, and type II CAH claims. This step identifies the FQHC, RHC, and type II CAH claims for FFS beneficiaries with dementia, retaining a limited set of variables needed for this step, including whether there was a dementia diagnosis on the claim, information on attending and rendering NPIs on the claims, and zip code of the facility. The code also merges the eligibility and demographic information from the 120 file onto these claims. Finally, the code creates the equivalent to the performing clinician NPI field in the carrier claims, using the attending NPI, if not missing, and otherwise the rendering NPI.
    - Step 4: Stack claims and roll up to the performing NPI level. In this step, the code stacks the carrier, Part D, and outpatient claims, and rolls up the claims to the performing NPI-level. In doing so, it creates the following variables: (1) the settings in which each NPI provided care to dementia patients (outpatient only; residential only; inpatient only; outpatient and residential; outpatient and inpatient; residential and inpatient; outpatient, inpatient, and residential; other settings only; or Part D only); (2) their practice zip code(s) (up to 10, ordered from billed to most often to billed to least often across all claims for the NPI); (3) the number of beneficiaries with dementia served (total number and number who

were DE/LIS);[8] and (4) the number of unique beneficiaries with dementia served who the NPI also diagnosed with dementia.

- 220_find_npis_mds_only.sas. This program implements the same broad steps described above for the 210 program, but it is limited to CY 2021 FFS claims and Part D drug claims for beneficiaries who were identified in the CY 2021 MDS only.

- 250_find_npis_allpts.sas. This program identifies all NPIs in carrier claims and FQHC, RHC, and type II CAH claims from the 100 program, and implements the broad steps described above for the 210 program to determine the settings in which NPIs provided care. This program uses all claims from these NPIs, not limited to claims for beneficiaries with dementia. Unlike the 210 program, however, this program excludes Part D claims. In fact, the output from this program is used in the 500 program to backfill setting information for NPIs who only showed up on Part D claims for beneficiaries with dementia (that is, we assume that these clinicians serve patients with dementia in the same settings in which they treat their patients without dementia).

- 310_find_ma_npis.sas. This program implements the same broad steps described above for the 210 program but uses the CY 2021 MA encounter data from the 101 program where records were flagged as having dementia diagnoses or not. However, to accommodate differences between the MA encounter records and FFS data, we made the following adjustments: (1) we limited the carrier and FQHC, RHC, and type II CAH encounter data to final action encounter records, excluding all chart review records; (2) we used header-level carrier data rather than line-level records as used in the 210 program (the encounter data headers have less missing data in the NPI fields and the place of service field compared with the line-level data, which had very high rates of missing data in these fields); and (3) to address many cases of missing NPI data in the encounter records, we used the rendering NPI on the carrier encounters (which is equivalent to the performing NPI on the FFS claims) where it was not missing and otherwise used the organizational NPI. Similarly, we used the attending NPI on outpatient records if not missing, else the rendering NPI if not missing, else the organizational NPI.

- 320_find_ma_npis_mds_only.sas. This program broadly implements the same steps as the 310 program but is limited to encounter records for beneficiaries who were identified in the MDS only and enrolled in MA during the year. Notably, this program does not use Part D claims data because the 220 program pulled Part D claims for all beneficiaries found in the MDS only and already identified all prescribing NPIs for these beneficiaries.

[8] The code also creates variables for the number of unique beneficiaries with dementia who were dually eligible, eligible for a low-income drug subsidy, and the number by various race and ethnicity categories, but we ultimately did not use these variables to develop the sample frame.

- 350_find_ma_npis_allpst.sas. This program mirrors the 250 program, which identifies the settings in which NPIs provide care for all beneficiaries. As noted in the 250 program, the output from this program is used in the 500 program to backfill setting data for NPIs who were only found in the Part D drug data.

- 410_merge_npis.sas. This program reads in the CY 2021 NPI-level output files from the 210 and 310 programs with information on number and types of FFS and MA beneficiaries with dementia served, respectively, as well as data on service settings. It then merges these two files on NPI to calculate the total number and type of beneficiaries with dementia served in CY 2021 (FFS plus MA), including the percentage of beneficiaries with dementia served by each NPI who are DE/LIS.[9] In the last step, it constructs a nine-category setting variable from the nine binary setting variables separately for FFS and MA, which is used downstream in the 500 program.

- 420_merge_npis_mds_only.sas. This program implements the same steps as the 410, but only for the NPIs who served beneficiaries with dementia identified in the MDS.

4. Identify the unique set of clinicians across CY 2021 and FY 2023 with relevant licensure and primary taxonomy and create variables needed for sampling

- 500_create_sample_frame.sas. This program implements the following steps to identify the initial set of NPIs for the sample frame:
    - Step 1: Merge primary taxonomy and priority address from NPPES onto NPI-level claims-based files. This step reads in the NPI-level file with all individual clinicians from the NPPES from the 020_npi program and creates a seven-category variable for clinician type that describes the licensure (MD/DO, NP, or PA) and specialty (primary care, neurology, and psychology) of clinicians sought for the survey. This variable is otherwise set to missing for NPIs with taxonomies that are not relevant to the survey, including several taxonomy codes for clinical nurse specialists that were initially flagged for inclusion but subsequently excluded. The program merges the NPPES data onto the CY 2021 NPI-level file from the 410 program and the FY 2023 file from the 210 program (in this file, the program also creates the nine-category setting variable that was created in the 410 program for clinicians in the CY 2021 data).
    - Step 2: Drop NPIs who did not have a primary taxonomy of relevance for the survey or were no longer active according to the NPPES.
    - Step 3: De-duplicate the list of NPIs. The program then identifies NPIs in the CY 2021 file that were also in the FY 2023 file and drops their CY 2021 records, so that each NPI is represented only once in the sample frame. The code also creates a four-category variable for whether the NPI was found in CY 2021 or FY 2023 data and whether the NPI was an MD/DO, NP, or PA.

[9] This program also calculates the percentage of beneficiaries with dementia served by FFS versus MA status and by race and ethnicity, but we did not use these variables for the sample frame.

- Step 4: Update the setting variable for NPIs identified as serving beneficiaries with dementia in Part D claims only. In this step, the code merges the NPI file above to the output files from the 250 and 350 programs—that is, NPI-level files with setting information based on claims and encounter records for all Medicare beneficiaries—and resets the setting variable, where possible, to use this information for NPIs with setting initially indicating "Part D prescriber." (Some NPIs had no carrier or outpatient claims, and their setting variable remained "Part D prescriber.")
- Step 5: Obtain additional geographic information for NPIs. The code reads in NPI-level data from PECOS with information on the state where each NPI practices and constructs a categorical variable for whether the NPI's PECOS data or NPPES priority address data indicate that they work outside of a U.S. state or district (for example, in a U.S. territory or elsewhere abroad). We used PECOS in addition to NPPES for this flag because the NPPES address data are not always current. We used this variable in the 700 program to exclude NPIs practicing outside of the United States. In this section of the program, the code also maps each NPI's most common practice zip code to its associated county using a publicly available crosswalk from the U.S. Census Bureau. Then, the code merges data on county-level MA and FFS enrollment onto the file from the publicly available Medicare Geographic Variations Database (GVDB) and calculates the percentage of Medicare beneficiaries enrolled in FFS in each county. We use this variable in the subsequent step.
- Step 6: Calculate the total number of beneficiaries with dementia served. In this step, the code calculates the total number of Medicare beneficiaries with dementia served by each NPI. For NPIs in the CY 2021 file only, this number is the actual number of FFS and MA beneficiaries with dementia served in CY 2021. For NPIs found in both FY 2023 and CY 2021 files, we assumed that their panel size of beneficiaries with dementia would be constant over time as well as the distribution of beneficiaries enrolled in FFS versus MA. Thus, for these NPIs, the code calculates their total panel size of Medicare beneficiaries with dementia as the number of MA beneficiaries served in CY 2021 plus the number of FFS beneficiaries served in FY 2023. For NPIs who were found in FY 2023 files only, the code predicts their MA panel size based on data from NPIs in both CY 2021 and FY 2023. Specifically, the code fits a linear regression on all NPIs in the FY 2023 file with the CY 2021 number of MA beneficiaries with dementia as the dependent variable. Predictor variables include: (1) FY 2023 number of FFS beneficiaries with dementia served, (2) clinician type, (3) nine-category setting variable, and (4) the percentage of Medicare beneficiaries in FFS in the county where the NPI practiced (per GVDB data). The code fits models for the overall number of CY 2021 MA beneficiaries with dementia served as well as number of CY 2021 MA beneficiaries who were DE/LIS and total CY 2021 MA beneficiaries who were Black or Hispanic/Latino. (As noted above, we ultimately did not use the race and ethnicity data to develop the sample frame.) The code then

truncates predicted counts outside of expected ranges (for example, less than zero or greater than the maximum number of observed MA beneficiaries). Specifically, the code sets predicted counts to zero if they were negative. If predicted counts were above the maximum observed number of patients across all NPIs, the code sets the predicted count to the 90th percentile of the number of MA beneficiaries. For predictions of the number of beneficiaries in MA who were DE/LIS that were above the maximum, the code sets the predicted counts to the minimum of the 90th percentile of the number of DE/LIS beneficiaries or the predicted number of MA beneficiaries (so that the predicted number of DE/LIS beneficiaries would not be larger than the predicted number of MA beneficiaries). For NPIs who were only in the FY 2023 file, the code calculates their panel size as the total number of FFS beneficiaries with dementia plus the predicted number of MA dementia patients in CY 2021.

- Step 7: Identify NPIs who served beneficiaries with dementia identified in the MDS only and who were not already in the set of NPIs above. This section of the code reads in the NPIs identified as having served beneficiaries with dementia who were identified in the CY 2021 MDS only. The code retains only those NPIs who were not found in the CY 2021 or FY 2023 data. It calculates their total panel size by summing the number of CY 2021 FFS and MA beneficiaries with dementia served and stacks these NPIs with the NPIs included in the steps above.
- Step 8: Create several alternative setting variables and save NPI-level output file. The program then creates a four-category mutually exclusive variable for setting based on the following logic: NPI was assigned to residential if they provided any services in any residential setting, else to outpatient if they provided any services in outpatient settings, else to Part D prescriber if they were found in Part D claims, else to all other settings. The code also creates an alternative measure that flips the order of the residential and outpatient categories and includes a category for providers serving beneficiaries in inpatient settings. The logic for the five-category mutually exclusive setting variable is the following: the NPI was assigned to outpatient if they provided any care in outpatient settings, else to residential if they provided any care in a residential setting, else to inpatient if they provided any care in an inpatient setting, else to Part D prescriber if they were found in Part D claims, else to all other settings. Finally, the code saves an output file of NPIs initially identified for the sample frame.

5. Finalize the sample frame after geocoding clinician's practice addresses to determine if they worked in rural regions and applying a final set of sample restrictions.

After running the programs above, we sent a file with all NPIs remaining at the end of the 500 program to IQVIA to obtain the NPIs' primary practice address. IQVIA searched two of its proprietary databases for addresses—one with primary practice address for NPIs affiliated with a health care organization and another for primary practice address for NPIs not in the affiliated

database. Thus, we received two files from IQVIA, one from each database. The steps below describe the final steps in developing the sample frame. Of note, we geocoded addresses (both from IQVIA and addresses from NPPES) using ArcGIS on a local computer. For this reason, there is no SAS code for the geocoding steps.

- 600_preprocess_iqvia.sas. This program reads in the two IQVIA data sets, stacks them, and saves the stacked file.

- 610_geocoded_diagnostics_iqvia.sas. This program takes the geocoded data from the manual ArcGIS processing and merges them with the output file from the 600 program. It then assesses the number of NPIs' addresses that ArcGIS was unable to geocode.

- 620_id_npi_tracts_iqvia.sas. This program assigns each geocoded address to the census tract where it is located using U.S. Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER)/line shapefile. The code runs the assignment process separately by state to speed up processing time. For a small number of addresses that did not map to a state federal processing standard publication (FIPS) code in the TIGER/line shapefile, but mapped to latitude and longitude, we rerun the assignment code using national (rather than state) data.

- 630_create_npi_rural_iqvia.sas. This program uses the output from the 620 program—that is, a file with each NPI and their census tract (if their address mapped)—to assign addresses to RUCA codes and to construct a binary rural variable based on the RUCA values. It uses the publicly available crosswalk data between 2010 census tract and RUCA code available from the U.S. Department of Agriculture (https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/documentation). The code flags a census tract as rural if its RUCA code falls between 7 and 10.

- 640_geocoded_diagnostics_nppes.sas. This program follows the same steps as in the 610 program except it is done for the priority address identified for each NPI from NPPES from the 020 program. We geocoded NPPES data to be used for any NPIs for whom IQVIA was unable to find in its database or whose IQVIA address ArcGIS was unable to geocode.

- 650_id_npi_rural_nppes.sas. This program follows the same steps as in the 620 program except it is done for the priority address identified for each NPI from NPPES from the 020 program.

- 660_create_npi_rural_nppes.sas. This program follows the same steps as in the 630 program except it is done for the priority address identified for each NPI from NPPES from the 020 program.

- 700_cc_same_frame.sas. This program creates the final sample frame for the community clinician survey. It merges on the rural flags based on IQVIA address (from the 630 program) and NPPES address (from the 660 program). It assigns each NPI to rural status based on the IQVIA rural flag, if populated, and otherwise the NPPES rural flag. The code also flags NPIs who appear to be working outside of the United States for exclusion. An NPI was considered to be working outside of the United States if (1) their IQVIA address did not geocode or it mapped to an undefined census tract *and* their PECOS or NPPES address data indicated they worked outside of a U.S. state or district (either in a U.S. territory or elsewhere overseas) or (2) their IQVIA address was in a U.S. territory. The code then applies the final set of restrictions—that is, it drops NPIs working outside of the United States and those working in inpatient settings only or all other settings only (that is, not outpatient, residential, or as a Part D prescriber). The code also masks cell sizes smaller than 11 by setting all such cells to 5. The code then limits the file to the variables needed and adds labels.

Appendix Table 1. ICD-10 diagnosis codes to identify patients with dementia

Dementia diagnosis codes	Description
Bynum-standard algorithm	
F01.50 – F01.51	Vascular dementia
F02.80 – F02.81	Dementia
F03.90-F03.91	Unspecified dementia
F04	Amnestic disorder
G30.0, G30.1, G30.8, G30.9	Alzheimer's disease
G31.01	Pick's disease
G31.09	Frontotemporal dementia
G31.83	Dementia with Lewy bodies
G31.1	Senile degeneration
G31.2	Degeneration of nervous system
R41.81	Age-related cognitive decline
New dementia-related ICD-10 codes effective as of October 1, 2022	
F01	Vascular dementia (non-billable code)
F015	Vascular dementia, unspecified severity (non-billable code)
F01.511	Vascular dementia, unspecified severity, with agitation
F01.518	Vascular dementia, unspecified severity, with other behavioral disturbance
F01.52	Vascular dementia, unspecified severity, with psychotic disturbance
F01.53	Vascular dementia, unspecified severity, with mood disturbance
F01.54	Vascular dementia, unspecified severity, with anxiety
F01.A	Vascular dementia, mild
F01.A0	Vascular dementia, mild, without behavioral disturbance, psychotic disturbance, mood disturbance, and anxiety
F01.A1	Vascular dementia, mild, with behavioral disturbance
F01.A11	Vascular dementia, mild, with agitation
F01.A18	Vascular dementia, mild, with other behavioral disturbance
F01.A2	Vascular dementia, mild, with psychotic disturbance
F01.A3	Vascular dementia, mild, with mood disturbance
F01.A4	Vascular dementia, mild, with anxiety
F01.B	Vascular dementia, moderate
F01.B0	Vascular dementia, moderate, without behavioral disturbance, psychotic disturbance, mood disturbance, and anxiety
F01.B1	Vascular dementia, moderate, with behavioral disturbance
F01.B11	Vascular dementia, moderate, with agitation
F01.B18	Vascular dementia, moderate, with other behavioral disturbance
F01.B2	Vascular dementia, moderate, with psychotic disturbance
F01.B3	Vascular dementia, moderate, with mood disturbance
F01.B4	Vascular dementia, moderate, with anxiety
F01.C	Vascular dementia, severe
F01.C0	Vascular dementia, severe, without behavioral disturbance, psychotic disturbance, mood disturbance, and anxiety

F01.C1	Vascular dementia, severe, with behavioral disturbance
F01.C11	Vascular dementia, severe, with agitation
F01.C18	Vascular dementia, severe, with other behavioral disturbance
F01.C2	Vascular dementia, severe, with psychotic disturbance
F01.C3	Vascular dementia, severe, with mood disturbance
F01.C4	Vascular dementia, severe, with anxiety
F02	Dementia in diseases classified elsewhere (non-billable code)
F028	Dementia in diseases classified elsewhere, unspecified severity
F02.811	Dementia in other diseases classified elsewhere, unspecified severity, with agitation
F02.818	Dementia in other diseases classified elsewhere, unspecified severity, with other behavioral disturbance
F02.82	Dementia in other diseases classified elsewhere, unspecified severity, with psychotic disturbance
F02.83	Dementia in other diseases classified elsewhere, unspecified severity, with mood disturbance
F02.84	Dementia in other diseases classified elsewhere, unspecified severity, with anxiety
F02.A	Dementia in other diseases classified elsewhere, mild
F02.A0	Dementia in other diseases classified elsewhere, mild, without behavioral disturbance, psychotic disturbance, mood disturbance, anxiety
F02.A1	Dementia in other diseases classified elsewhere, mild, with behavioral disturbance
F02.A11	Dementia in other diseases classified elsewhere, mild, with agitation
F02.A18	Dementia in other diseases classified elsewhere, mild, with other behavioral disturbance
F02.A2	Dementia in other diseases classified elsewhere, mild, with psychotic disturbance
F02.A3	Dementia in other diseases classified elsewhere, mild, with mood disturbance
F02.A4	Dementia in other diseases classified elsewhere, mild, with anxiety
F02.B	Dementia in other diseases classified elsewhere, moderate
F02B0	Dementia in other diseases classified elsewhere without behavioral disturbance, mood disturbance, and anxiety
F02.B1	Dementia in other diseases classified elsewhere, moderate, with behavioral disturbance
F02.B11	Dementia in other diseases classified elsewhere, moderate, with agitation
F02.B18	Dementia in other diseases classified elsewhere, moderate, with other behavioral disturbance
F02.B2	Dementia in other diseases classified elsewhere, moderate, with psychotic disturbance
F02.B3	Dementia in other diseases classified elsewhere, moderate, with mood disturbance
F02.B4	Dementia in other diseases classified elsewhere, moderate, with anxiety
F02.C	Dementia in other diseases classified elsewhere, severe
F02.C0	Vascular in other diseases classified elsewhere, without behavioral disturbance, psychotic disturbance, mood disturbance, and anxiety

F02.C1	Dementia in other diseases classified elsewhere, severe, with behavioral disturbance
F02.C11	Dementia in other diseases classified elsewhere, severe, with agitation
F02.C18	Dementia in other diseases classified elsewhere, severe, with other behavioral disturbance
F02.C2	Dementia in other diseases classified elsewhere, severe, with psychotic disturbance
F02.C3	Dementia in other diseases classified elsewhere, severe, with mood disturbance
F02.C4	Dementia in other diseases classified elsewhere, severe, with anxiety
F03	Unspecified dementia (non-billable code)
F039	Unspecified dementia (also non-billable)
F03.911	Unspecified dementia, unspecified severity, with agitation
F03.918	Unspecified dementia, unspecified severity, with other behavioral disturbance
F03.92	Unspecified dementia, unspecified severity, with psychotic disturbance
F0393	Unspecified dementia, unspecified severity, with mood disturbance
F03.94	Unspecified dementia, unspecified severity, with anxiety
F03.A	Unspecified dementia, mild
F03.A0	Unspecified dementia, mild, without behavioral disturbance, psychotic disturbance, mood disturbance, anxiety
F03.A1	Unspecified dementia, mild, with behavioral disturbance
F03.A11	Unspecified dementia, mild, with agitation
F03.A18	Unspecified dementia, mild, with other behavioral disturbance
F03.A2	Unspecified dementia, mild, with psychotic disturbance
F03.A3	Unspecified dementia, mild, with mood disturbance
F03.A4	Unspecified dementia, mild, with anxiety
F03.B	Unspecified dementia, moderate
F03.B0	Unspecified dementia, moderate, without behavioral disturbance, psychotic disturbance, mood disturbance, anxiety
F03.B1	Unspecified dementia, moderate, with behavioral disturbance
F03.B11	Unspecified dementia, moderate, with agitation
F03.B18	Unspecified dementia, moderate, with other behavioral disturbance
F03.B2	Unspecified dementia, moderate, with psychotic disturbance
F03.B3	Unspecified dementia, moderate, with mood disturbance
F03.B4	Unspecified dementia, moderate, with anxiety
F03.C	Unspecified dementia, severe
F03.C0	Unspecified dementia, severe, without behavioral disturbance, psychotic disturbance, mood disturbance, anxiety
F03.C1	Unspecified dementia, severe, with behavioral disturbance
F03.C11	Unspecified dementia, severe, with agitation
F03.C18	Unspecified dementia, severe, with other behavioral disturbance
F03.C2	Unspecified dementia, severe, with psychotic disturbance
F03.C3	Unspecified dementia, severe, with mood disturbance
F03.C4	Unspecified dementia, severe, with anxiety
G30	Alzheimer's disease (non-billable code)

G310	Frontotemporal dementia (non-billable code)

Notes:
(a) Online supplementary materials from Grodstein F., C.-H. Chang, A. W. Capuano, M.C. Power, D.X. Marquez, L.L. Barnes, et al. "Identification of Dementia in Recent Medicare Claims Data, Compared with Rigorous Clinical Assessments." *The Journals of Gerontology, Series A,* vol. 77, no. 6, 2021, pp. 1272–1278. https://doi.org/10.1093/gerona/glab377.
(b) Per Grodstein et al. (2021): "These ICD codes are searched in the following claims files:
1) Any inpatient or skilled nursing facility claim
2) Home Health Agency: Any claim
3) Hospice: Any claim
4) Hospital Outpatient File (HOF) for outpatient medical services: Only claims from Rural Health Clinics, Federally Qualified Health Centers, and Critical Access Hospitals-Payment Option II
5) Carrier (Provider) File for services from physicians and other health care providers: Any claim, with the condition that there are two or more qualifying Carrier or HOF claims at least 7 days apart."

Appendix Table 2. Setting assignment based on codes for place of service (carrier data) and revenue center (outpatient data from FQHCs, RHCs, and type II CAHs)

Setting	Place of service values (carrier claims)	Revenue center codes (outpatient claims)
Outpatient	05 = Indian health service (free-standing facility) 07 = Tribal 638 (free-standing facility) 10 = Telehealth provided in patient's home 11 = Office 17 = Walk-in retail health clinic 19 = Off campus (outpatient hospital) 20 = Urgent care facility 22 = On campus (outpatient hospital) 49 = Independent clinic 50 = Federally Qualified Health Center 62 = Comprehensive outpatient rehabilitation facility 71 = Public health clinic 72 = Rural health clinic	All other values on claims from FQHCs, RHCs, and type II CAHs not used below
Inpatient	06 = Indian health service (provider-based facility) 08 = Tribal 638 provider-based facility 21 = Inpatient hospital 61 = Comprehensive inpatient rehabilitation facility	n.a.
Residential	13 = Assisted living facility 14 = Group home 31 = Skilled nursing facility 32 = Nursing facility 33 = Custodial care facility 34 = Hospice 54 = Intermediate care / intellectual disabilities facility	0524 = free-standing clinic – visit by RHC/FQHC practitioner to a member in a covered Part A stay at the SNF 0525 = free-standing clinic – visit by RHC/FQHC practitioner to a member in a SNF (not in a covered Part A stay) or NF or ICF MR or other residential facility
All other settings	All other values not used above (e.g., 23 = ER [hospital], 53 = community mental health center)	0522 = Free-standing clinic – home visit by RHC/FQHC practitioner 0527 = Free-standing clinic – RHC/FQHC visiting nurse service(s) to a member's home when in a home health shortage area

		0528 = Free standing clinic – visit by RHC/FQHC practitioner to other non-RHC/FQHC site (e.g., scene of accident)

Notes:
(a) The codes for place of service and revenue center apply to both fee-for-service claims and Medicare Advantage encounter records.
(b) We assigned clinicians who appeared in Part D drug claims only to the "Part D prescriber" setting.
CAH = critical access hospital; FQHC = Federally qualified health center; n.a. = not applicable; NF = nursing facility; ICF MR = intermediate care facility for individuals with intellectual disabilities; RHC = rural health center; SNF = skilled nursing facility.

Appendix Table 3. Taxonomy codes used to identify clinicians with primary care, neurology, or psychiatry specialties

Taxonomy code	Definition	Specialty	Licensure
207LH0002X	Allopathic & Osteopathic Physicians/Anesthesiology/Hospice and Palliative Medicine	Primary care	MD/DO
207PH0002X	Allopathic & Osteopathic Physicians/Surgery/Hospice and Palliative Medicine, Emergency Medicine	Primary care	MD/DO
207Q00000X	Allopathic & Osteopathic Physicians/Family Medicine	Primary care	MD/DO
207QA0000X	Allopathic & Osteopathic Physicians/Family Medicine, Adolescent Medicine	Primary care	MD/DO
207QA0401X	Allopathic & Osteopathic Physicians/Family Medicine, Addiction Medicine	Primary care	MD/DO
207QA0505X	Allopathic & Osteopathic Physicians/Family Medicine, Adult Medicine	Primary care	MD/DO
207QG0300X	Allopathic & Osteopathic Physicians/Family Medicine, Geriatric Medicine	Primary care	MD/DO
207QH0002X	Allopathic & Osteopathic Physicians/Family Medicine, Hospice and Palliative Medicine	Primary care	MD/DO
207QS0010X	Allopathic & Osteopathic Physicians/Family Medicine, Sports Medicine	Primary care	MD/DO
207QS1201X	Allopathic & Osteopathic Physicians/Family Medicine, Sleep Medicine	Primary care	MD/DO
207R00000X	Allopathic & Osteopathic Physicians/Internal Medicine	Primary care	MD/DO
207RA0000X	Allopathic & Osteopathic Physicians/Internal Medicine, Adolescent Medicine	Primary care	MD/DO
207RA0002X	Allopathic & Osteopathic Physicians/Internal Medicine/Adult Congenital Heart Disease	Primary care	MD/DO
207RA0401X	Allopathic & Osteopathic Physicians/Internal Medicine, Addiction Medicine	Primary care	MD/DO
207RG0300X	Allopathic & Osteopathic Physicians/Internal Medicine, Geriatric Medicine	Primary care	MD/DO
207RH0002X	Allopathic & Osteopathic Physicians/Internal Medicine, Hospice and Palliative Medicine	Primary care	MD/DO
207RS0010X	Allopathic & Osteopathic Physicians/Internal Medicine, Sports Medicine	Primary care	MD/DO
207RS0012X	Allopathic & Osteopathic Physicians/Internal Medicine, Sleep Medicine	Primary care	MD/DO
207VH0002X	Allopathic & Osteopathic Physicians/Obstetrics & Gynecology, Hospice and Palliative Medicine	Primary care	MD/DO
2080H0002X	Allopathic & Osteopathic Physicians/Pediatrics, Hospice and Palliative Medicine	Primary care	MD/DO
2081H0002X	Allopathic & Osteopathic Physicians/Physical Medicine & Rehabilitation, Hospice and Palliative Medicine	Primary care	MD/DO
2083P0901X	Allopathic & Osteopathic Physicians/Preventive Medicine, Public Health & General Preventive Medicine	Primary care	MD/DO

2084H0002X	Allopathic & Osteopathic Physicians/Surgery/Hospice and Palliative Medicine, Neuropsychiatry	Primary care	MD/DO
2085H0002X	Hospice & Palliative Medicine	Primary care	MD/DO
2086H0002X	Allopathic & Osteopathic Physicians/Surgery/Hospice and Palliative Medicine	Primary care	MD/DO
208D00000X	Allopathic & Osteopathic Physicians/General Practice	Primary care	MD/DO
363A00000X	Physician Assistants & Advanced Practice Nursing Providers/Physician Assistant	Primary care	PA
363AM0700X	Physician Assistants & Advanced Practice Nursing Providers/Physician Assistant, Medical	Primary care	PA
363L00000X	Physician Assistants & Advanced Practice Nursing Providers/Nurse Practitioner	Primary care	NP
363LA2200X	Physician Assistants & Advanced Practice Nursing Providers/Nurse Practitioner, Adult Health	Primary care	NP
363LC1500X	Physician Assistants & Advanced Practice Nursing Providers/Nurse Practitioner, Community Health	Primary care	NP
363LF0000X	Physician Assistants & Advanced Practice Nursing Providers/Nurse Practitioner, Family	Primary care	NP
363LG0600X	Physician Assistants & Advanced Practice Nursing Providers/Nurse Practitioner, Gerontology	Primary care	NP
363LP2300X	Physician Assistants & Advanced Practice Nursing Providers/Nurse Practitioner, Primary Care	Primary care	NP
2084A0401X	Allopathic & Osteopathic Physicians/Psychiatry & Neurology, Addiction Medicine	Psychiatry	MD/DO
2084B0002X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Bariatric Medicine	Psychiatry	MD/DO
2084B0040X	Allopathic & Osteopathic Physicians/Psychiatry & Neurology/Behavioral Neurology & Neuropsychiatry	Psychiatry	MD/DO
2084F0202X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Forensic Psychiatry	Psychiatry	MD/DO
2084P0015X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Psychosomatic Medicine	Psychiatry	MD/DO
2084P0301X	Allopathic & Osteopathic Physicians/Psychiatry & Neurology/Respiratory, Developmental, Rehabilitative and Restorative Service, Brain Injury Medicine	Psychiatry	MD/DO
2084P0800X	Allopathic & Osteopathic Physicians/Psychiatry	Psychiatry	MD/DO
2084P0802X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Addiction Psychiatry	Psychiatry	MD/DO
2084P0804X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Child & Adolescent Psychiatry	Psychiatry	MD/DO
2084P0805X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Geriatric Psychiatry	Psychiatry	MD/DO
2084P2900X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Pain Medicine	Psychiatry	MD/DO
2084S0010X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Sports Medicine	Psychiatry	MD/DO

2084S0012X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Sleep Medicine	Psychiatry	MD/DO
363LP0808X	Physician Assistants & Advanced Practice Nursing Providers/Nurse Practitioner, Psychiatric/Mental Health	Psychiatry	NP
2084A2900X	Allopathic & Osteopathic Physicians/Psychiatry & Neurology/Neurocritical Care	Neurology	MD/DO
2084D0003X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Diagnostic Neuroimaging	Neurology	MD/DO
2084E0001X	Allopathic & Osteopathic Physicians/Psychiatry & Neurology/Epilepsy	Neurology	MD/DO
2084N0008X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Neuromuscular Medicine	Neurology	MD/DO
2084N0400X	Allopathic & Osteopathic Physicians/Psychiatry and Neurology, Neurology	Neurology	MD/DO
2084N0402X	Allopathic & Osteopathic Physicians/Psychiatry and Neurology, Neurology with Special Qualifications in Child Neurology	Neurology	MD/DO
2084N0600X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Clinical Neurophysiology	Neurology	MD/DO
2084P0005X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Neurodevelopmental Disabilities	Neurology	MD/DO
2084V0102X	Allopathic & Osteopathic Physicians/ Psychiatry & Neurology, Vascular Neurology	Neurology	MD/DO

Appendix Table 4. Sample frame counts and inclusion and exclusion criteria

	Clinician counts	
	FY 2023 FFS data	**CY 2021 FFS and MA data**
NPIs with a claim from a dementia patient	1,318,196	1,647,563
Exclusions:		
Organizational NPIs	61,763	171,484
NPIs who do not have a primary taxonomy of primary care, psychology, or neurology	716,999	888,047
NPIs who are no longer active by April 2024	2,202	7,898
Subtotal: NPIs with relevant taxonomy and are active in April 2024 per NPPES	537,232	580,134
	Total number of clinicians	
Number of NPIs from the 2021 and FY 2023 files (this is the sum of NPIs in row above). This is not a de-duplicated count—that is, the same NPI might be in both files.	1,117,366	
Additions:		
NPIs who cared for beneficiaries with dementia identified in the 2021 MDS but not 2021 claims or encounter records, and where the NPI has a relevant primary taxonomy, is still active in April 2024, and was not found in the set of NPIs above	308	
Exclusions:		
Duplicate NPIs from the 2021 file who were also in the FY 2023 file or NPIs who were FFS only in 2021 and no longer in FFS in 2023 and presumably retired[a]	553,135	
NPIs working in inpatient settings only	37,867	
NPIs working in "all other settings" only	31,649	

NPIs working outside of the United States	2,840
Final set of NPIs	**492,183**

[a] The NPIs from 2021 retained in the sample frame were those who served beneficiaries with dementia enrolled in MA only. We assumed that these clinicians continued to serve MA-enrolled beneficiaries (for example, clinicians working in health systems such as Kaiser Permanente or Geisinger).

CY 2021 = calendar year 2021; FFS = fee-for-service; FY 2023 = federal fiscal year 2023 (October 1, 2022 – September 30, 2023); MA = Medicare Advantage; MDS = Minimum Data Set; NPI = National Provider Identifier; NPPES = National Plan and Provider Enumeration System.