NEXT STEPs: Nursing Home Data Sources Compendium

Data Source	Description	Key Variables	Access (Public/Restricted) & Process	HIPAA-protected (yes/no)	Smallest Unit of Analysis (e.g., resident, facility)	Other Possible Units of Analysis	Linkable & At What Unit(s)	Benefits	Limitations	Use Case Examples	Cost (0=free, $=low, $$=moderate, $$$=high)	URL to Access Data
Agency for Healthcare Research and Quality (AHRQ) Nursing Home Survey on Patient Safety Culture	A standardized, validated survey developed by AHRQ to assess the culture of patient safety within NHs, focusing on staff perspectives about communication, teamwork, management support, and resident safety practices.	Staff responses to items across multiple safety domains (e.g., communication openness, non-punitive response to mistakes, staffing, training, teamwork, handoffs, and transitions); facility characteristics.	Public; survey instrument and aggregate results are freely available on the AHRQ website. NHs can administer the survey internally or submit data voluntarily to AHRQ's Comparative Database. Access to raw individual-level response data from the Comparative Database requires a data use agreement.	No (data are de-identified and do not contain resident-level health information)	Staff respondent	Facility, Unit within facility	Not linkable to other external datasets; facility-level data may be linked internally with staffing or quality data by the facility itself	Provides a standardized tool for assessing organizational safety culture; facilitates benchmarking against national norms; identifies areas for improvement in safety-related communication and practice.	Data submission is voluntary; limited external comparability if facilities do not participate in AHRQ's Comparative Database; no direct link to resident outcomes or clinical quality measures.	Assessing staff perceptions of safety culture, evaluating the impact of training or policy changes on perceived safety climate, guiding quality improvement initiatives in NHs.	0	https://www.ahrq.gov/sops/surveys/nursing-home/index.html
Health and Retirement Study (HRS) - Linked	HRS survey data linked to Medicare claims (Parts A, B, C, D), Medicaid data, and Minimum Data Set (MDS) assessments, enabling detailed longitudinal analyses of healthcare utilization and outcomes in the context of rich sociodemographic, cognitive, and end-of-life preference data.	All unlinked HRS variables plus: Medicare claims (hospital, SNF, hospice, outpatient, Part D), Medicaid claims, MDS clinical assessments, dates of NH stays, diagnosis and procedure codes, and mortality indicators.	Restricted; requires HRS and CMS data use agreements, IRB approval, and submission of a restricted data request through the HRS website.	Yes	Individual respondent	Resident, Episode of Care, Facility (limited), Household, Survey Wave	Yes, resident-level; technically facility-level via MDS but not well-suited for facility-level inference	Combines rich psychosocial, cognitive, and preference data with healthcare utilization and NH clinical assessments; supports end-of-life, dementia, and long-term services research in nationally representative older adults.	Restricted to HRS participants who consented to linkage; not representative of NHs; not designed for facility-level analysis; linkage periods vary by data source.	Evaluating life-sustaining treatment use and treatment concordance among residents with dementia in their last 90 days of life; identifying associations between advance care planning and intensity of NH or hospital care; stratifying care patterns by NH use.	$	https://hrs.isr.umich.edu/data-products/restricted-data
Health and Retirement Study (HRS) - Unlinked	A longitudinal panel survey of U.S. adults over age 50, designed to examine aging, retirement, health, and economic well-being. The unlinked HRS includes interview data, physical and psychosocial measures, and self-reported health and healthcare use.	Demographics, income and wealth, functional and cognitive status, self-reported chronic conditions, insurance coverage, caregiving, residential transitions, end-of-life planning.	Public; most data are freely available upon registration via the HRS website. Sensitive files (e.g., geographic detail) require a restricted data application.	No (unlinked survey data are not considered HIPAA-protected)	Individual respondent	Household, Cohort, Survey Wave	Yes, resident-level, facility-level	Rich, nationally representative longitudinal survey with psychosocial and economic depth; captures preferences, caregiving dynamics, and transitions relevant to NH and long-term care planning.	Does not contain clinical or claims data; relies on self-report; NH residents underrepresented; not suitable for analyzing care delivery or outcomes.	Examining predictors of NH entry, financial stress and caregiving burden, or advance care planning among older adults at risk for institutional care.	0	https://hrs.isr.umich.edu/data-products
National Dementia Workforce Survey (NDWS)	A nationally representative survey of the dementia care workforce in long-term care settings, including NHs and assisted living. Conducted to understand training, attitudes, work environment, and care practices of direct care workers and licensed staff.	Staff role and training, dementia-specific care experience, attitudes toward dementia, staffing and workload, burnout, care strategies.	Restricted; access may be available via request to the survey team or affiliated research institutions. No public use files currently posted.	No (staff self-report survey; does not include resident health data)	Individual staff member	Facility, Occupation, Region	Possibly linkable at facility level, but linkage is limited and requires permission	Unique insights into the experiences, perspectives, and practices of the NH dementia care workforce; enables study of workforce preparedness and training needs.	Not yet longitudinal; self-reported data may be subject to bias; access to data is limited and not publicly posted.	Evaluating staff preparedness to care for residents with dementia; assessing burnout, attitudes, or care strategies in NH dementia care workforce.	0	https://www.ndws.org/surveys-and-data/how-to-access-data
National Long Term Care Survey (NLTCS) - Linked	A longitudinal survey of Medicare beneficiaries aged 65+ linked to Medicare claims data (Parts A, B, D), Medicaid claims, and/or MDS assessments. Enables detailed analyses of long-term care utilization and outcomes in relation to functional and cognitive status.	All NLTCS survey variables (e.g., ADLs/IADLs, cognitive status, caregiver burden) plus: Medicare claims (diagnoses, procedures, payments), Medicaid claims, MDS assessments (functional/cognitive status, services received), mortality data.	Restricted; requires application and data use agreements via Duke University or CMS (ResDAC), depending on the linked file. Some may be available through the NIA LINKAGE Program.	Yes	Claim	Resident, Episode of Care, Facility (limited), Survey Wave, Region	Yes, resident-level; technically facility-level via MDS but not well-suited for facility-level inference	Combines detailed survey data with claims-based outcomes; supports research on transitions of care, disability, and service use over time.	Linkage limited to survey respondents who consented; data are dated (last survey in 2004); not fully representative of current NH population; limited facility-level detail.	Studying transitions into and out of NHs, evaluating disability-related care use and spending, analyzing trajectories of functional decline and service utilization.	$	https://www.icpsr.umich.edu/web/NACDA/studies/9681

Data Source	Description	Key Variables	Access (Public/Restricted) & Process	HIPAA-protected (yes/no)	Smallest Unit of Analysis (e.g., resident, facility)	Other Possible Units of Analysis	Linkable & At What Unit(s)	Benefits	Limitations	Use Case Examples	Cost (0=free, $=low, $$=moderate, $$$=high)	URL to Access Data
National Long Term Care Survey (NLTCS) - Unlinked	A longitudinal survey of Medicare beneficiaries aged 65+ designed to track changes in health and functioning related to long-term care needs. The survey was conducted in 1982, 1984, 1989, 1994, 1999, and 2004. Focuses on disability, health status, family support, and care arrangements.	Functional status (ADLs/IADLs), chronic conditions, demographics, living arrangements, use of formal and informal care, caregiving burden, service use, cognitive status.	Public use files available via Duke University and NIA. Sensitive or restricted versions (with geographic or detailed clinical variables) require an application.	No for public use files; Yes for restricted files.	Individual respondent	Household, Survey Wave, Region	Yes, resident-level; potentially facility-level via MDS, but limited inference	Rich longitudinal information on health and disability among older adults; captures transitions between home and institutional settings; nationally representative of older Medicare beneficiaries.	Last fielded in 2004; data are dated; no resident identifiers or consistent facility-level information; lacks administrative claims or clinical outcomes.	Examining long-term disability trajectories, informal caregiving patterns, or predictors of NH entry among older adults with functional impairments.	0	https://www.icpsr.umich.edu/web/NACDA/studies/9681
National Nursing Home Survey (NNHS) / National Study of Long-Term Care Providers (NSLTCP)	NNHS (conducted 1973–2004) was a nationally representative survey of NHs, their residents, staff, and characteristics. In 2012, NNHS was replaced by the NSLTCP, which continues the collection of long-term care provider data, including residential care communities and adult day services, but no longer includes resident-level data for NHs. Together, they provide complementary snapshots of the NH sector across different time periods and methodologies.	NNHS: Resident demographics, diagnoses, medication use, care practices, facility characteristics, staffing levels. NSLTCP: Provider characteristics, staffing, services offered, organizational structure, ownership, and select aggregate resident characteristics.	NNHS: Public-use and restricted-use files available via NCHS and RDC (Research Data Center). NSLTCP: Publicly accessible summary tables and microdata (depending on component); restricted files via NCHS RDC.	No for public use files; Yes for restricted files.	NNHS: Resident NSLTCP: Facility	Staff, Provider Type, Region	Limited; no consistent resident identifiers or linkage to claims data. Some facility-level linkage possible for NSLTCP with external datasets (e.g., using provider characteristics or location identifiers).	NNHS provides rare nationally representative resident-level survey data on NH care in earlier decades. NSLTCP provides ongoing national data on the structure and characteristics of long-term care providers, including residential care.	NNHS data are outdated (last collected in 2004); NSLTCP no longer includes resident-level data for NHs; linking to other datasets is difficult or impossible; limited clinical detail compared to EHR or claims data.	NNHS: Examining medication use patterns (e.g., PPI prescribing without clear indication) among NH residents. NSLTCP: Describing variation in hospitalization and readmission rates across residential care settings and identifying organizational characteristics that reduce risk.	0 / $	https://www.cdc.gov/rdc/restricted-nchs-variables/nnhs-nnas.html; https://www.cdc.gov/nchs/npals/index.htm
National Post-acute and Long-term Care Study (NPALS)	Nationally representative, periodic cross-sectional survey conducted by the National Center for Health Statistics (NCHS) to assess organizational characteristics, services, staffing, and resident demographics of long-term care providers including nursing homes, assisted living, and home health agencies. Replaced the NSLTCP in 2020.	Facility ownership and size, staffing levels, resident capacity and demographics, services provided, end-of-life care practices, admission and discharge practices, infection control policies (during COVID-19 years).	Public-use datasets are available from NCHS. Restricted-use files with more detailed variables and geographic identifiers are available through the NCHS Research Data Center (RDC) with an approved proposal.	No for public-use data; Yes for restricted-use data.	Facility	Geographic region, Provider type	No direct linkage to other datasets; may be linkable by geographic unit with RDC approval	Publicly available and nationally representative; captures policy-relevant characteristics of post-acute and long-term care providers; includes consistent questions across waves to assess trends.	Cross-sectional; self-reported data; no resident-level detail; not designed for longitudinal tracking or linkage with clinical outcomes.	Analyzing national trends in NH staffing and services; assessing changes in infection control or end-of-life care policies; comparing facility characteristics across provider types.	0 / $	https://www.cdc.gov/nchs/npals/index.htm
Nursing Home Component of the National Health and Aging Trends Study (NHATS) - Linked	NHATS survey data linked to Medicare claims (Parts A, B, D), Medicaid claims, MDS assessments, and provider files. Enables robust analyses of healthcare utilization, service transitions, and clinical outcomes among NH residents.	All NHATS survey variables plus: Medicare/Medicaid claims (diagnoses, procedures, costs), MDS clinical assessments, provider data, mortality indicators.	Restricted; requires NHATS, CMS, and possibly ResDAC data use agreements and approvals; IRB approval required. Some files accessible via the NIA LINKAGE Program.	Yes	Claim	Resident, Episode of Care, Facility (limited), Survey Wave	Yes, resident-level; technically facility-level via MDS and provider files, but not ideal for facility inference	Enables integrated analysis of resident function, preferences, and cognition with healthcare utilization, transitions, and outcomes; supports research on end-of-life, dementia, and aging in place.	Restricted to NHATS participants who consented to linkage; NH sample may underrepresent long-stay residents; linkage periods vary across data sources.	Evaluating end-of-life care, care transitions, cognitive and functional decline, or hospitalizations among NH residents using linked clinical and survey data.	$	https://www.nhats.org/researcher/data-access
Nursing Home Component of the National Health and Aging Trends Study (NHATS) - Unlinked	A nationally representative longitudinal survey of Medicare beneficiaries aged 65+ that includes in-person interviews, cognitive testing, physical performance measures, and facility questionnaires for NH residents.	Resident demographics, physical and cognitive functioning, health conditions, caregiving, NH environment, staffing, end-of-life care preferences.	Public use files available from NHATS website; restricted files (e.g., facility characteristics, geographic detail) require application and approval.	No for public use files; Yes for restricted files	Individual respondent	Facility, Survey Wave	Yes, resident-level; technically facility-level via NH Facility Questionnaire, but not recommended for facility-level inference	Captures rich, multidimensional data from NH residents, including cognitive/physical assessments and self-report data; supports person-centered aging and care research.	Facility-level analysis limited; data not suitable for detailed health service utilization or provider-level comparisons; sample may underrepresent long-stay NH residents.	Assessing quality of life, cognition, function, caregiving burden, and preferences among NH residents; analyzing resident characteristics over time.	0	https://www.nhats.org/researcher/data-access

Data Source	Description	Key Variables	Access (Public/Restricted) & Process	HIPAA-protected (yes/no)	Smallest Unit of Analysis (e.g., resident, facility)	Other Possible Units of Analysis	Linkable & At What Unit(s)	Benefits	Limitations	Use Case Examples	Cost (0=free, $=low, $$=moderate, $$$=high)	URL to Access Data
Brown University's LTCFocus	Aggregated dataset that compiles nursing home resident characteristics, facility attributes, and regional market information from multiple national sources (e.g., MDS, OSCAR/CASPER).	Facility-level resident characteristics (e.g., demographics, functional status, cognitive status), facility characteristics (e.g., size, ownership, staffing levels), occupancy rates, regional healthcare market characteristics.	Public; available via Brown University's LTCFocus website. Users can download data directly after creating a free account.	No	Facility	County, Market, State	Yes, facility-level and geographic-level	User-friendly aggregated data, ideal for market-level analyses; integrates multiple CMS datasets; regularly updated; easily accessible.	Lacks individual-level resident data; does not provide detailed clinical variables beyond MDS aggregates; no real-time updates (updated annually or bi-annually); some missing data due to small cell suppression requirements; numerical data coded as text and requires transformation.	Studying NH market characteristics, evaluating staffing patterns and occupancy trends, assessing regional variations in NH resident composition and care quality.	0	https://ltcfocus.org/
CDC's National Healthcare Safety Network (NHSN) Long Term Care Facility (LTCF) COVID-19 Module	Mandatory reporting system initiated by CDC in May 2020 to collect weekly data from LTCFs on COVID-19 cases, deaths, testing, staffing shortages, and PPE availability. Used to support pandemic monitoring and response.	Weekly counts of resident and staff COVID-19 cases and deaths, number of new admissions, vaccination rates, PPE availability, staffing shortages, and mitigation strategies.	Publicly available data files are posted by CMS; facility-level data are downloadable. Additional detail may be available through CDC data requests.	No	Facility	State, Region, Chain (via name), Time (week)	Yes, facility-level	Near real-time, comprehensive national surveillance of COVID-19 in LTCFs; critical for tracking pandemic trends, resource needs, and facility responses.	Self-reported data; potential for underreporting or inconsistent definitions across facilities; limited to pandemic-related variables.	Monitoring infection trends and outbreak burden in NHs; evaluating impact of staffing shortages or PPE supply on COVID-19 outcomes; assessing vaccination rates.	0%	https://data.cms.gov/covid-19/covid-19-nursing-home-data
Centers for Medicare & Medicaid Services Payroll-Based Journal (CMS PBJ)	Federally mandated data capturing detailed staffing levels and hours worked by nursing home direct-care staff. Facilities submit data quarterly, ensuring standardized national reporting.	Staffing hours per resident per day for registered nurses, licensed practical nurses, certified nursing assistants, and other direct-care staff; staff turnover rates; facility-level staffing characteristics.	Public; data can be accessed via the CMS website without special permissions.	No	Facility	Market, State	Yes, facility-level	Reliable, standardized reporting of staffing levels across U.S. NHs; allows monitoring of compliance with federal staffing requirements; supports detailed workforce analysis.	Does not include detailed clinical outcomes or individual-level staffing assignments; relies on self-reporting by facilities, potentially leading to inaccuracies; lacks resident-level granularity.	Evaluating relationships between staffing levels, turnover, and NH quality; analyzing staffing adequacy and compliance with regulatory staffing guidelines.	0	https://data.cms.gov/search?keywords=PBJ&sort=Relevancy
Electronic Health Record (EHR) Data from Major Vendors (e.g., PointClickCare, MatrixCare, Netsmart) or NH Chain	Data from proprietary EHR systems used by nursing homes for clinical and administrative management.	Diagnoses, vital signs, progress notes, medication orders, care plans, laboratory results.	Restricted; requires agreements with vendors (PointClickCare, MatrixCare, etc.) or with NH chain.	Yes	Clinical measurement (e.g., medication administration, laboratory measures, vital signs)	Resident, Facility	Yes, resident-level	Rich clinical data in real-time, used for direct patient care; information absent from other datasets is available (e.g., laboratory values).	Vendor-specific formats create challenges for standardization and research linkage; vendors may restrict certain data to protect resident privacy. Data are comparatively expensive.	Analyzing care pathways, medication prescribing patterns, and resident outcomes.	$$$	https://pointclickcare.com/industry-solutions/life-sciences/real-world-data/ ; https://www.matrixcare.com/; https://www.ntst.com/
IQVIA LTC-LRx Pharmacy Data Product	Commercially available prescription dispensing data covering medications provided to residents in LTC facilities across the U.S., based on pharmacy transactions.	Drug name, dosage, quantity, date dispensed, prescriber identifier, facility identifier, days' supply.	Restricted; requires purchase and licensing through IQVIA, typically with tailored data use agreements.	Yes	Claim	Resident, Prescriber, Facility, Pharmacy	Yes, resident-level (with appropriate linkage); potentially facility-level	Broad national scope; includes data from multiple pharmacy providers across LTC settings; captures real-world prescribing patterns and medication use.	No public access; limited clinical information or outcome data; costly to access; exact facility or resident linkage requires customization; may not be possible to stratum on type of LTC facility (e.g., assisted living versus nursing home)	Describing trends in prescribing practices in LTC settings; evaluating use of specific drug classes (e.g., antipsychotics, opioids) over time; benchmarking medication utilization across facilities.	$$$	https://www.iqvia.com/insights/the-iqvia-institute/available-iqvia-data
Long-Term Care Data Cooperative (LTCDC) - Linked	LTCDC data linked to nationwide data from the Centers for Medicare and Medicaid Services (CMS) (claims, encounter and enrollment records, assessments) through the NIA LINKAGE Program to provide additional information before, during and after a NH stay.	All information in the unlinked LTCDC data, plus diagnoses, billing codes, and other information available in research identifiable datafiles from CMS (Parts A/B/C/D, Medicaid, and more).	Restricted; requires LTCDC and NIA approvals and data use agreements.	Yes	Clinical measurement (e.g., medication administration, laboratory measure, vital sign)	Resident, Facility	Yes, resident-level	Provides a more complete understanding of resident healthcare utilization before, during, and after the nursing home stay.	Similar to LTCDC-unlinked; Additional approvals required.	Examining care transitions, use of services before, during and after a nursing home stay.	0 for standard access; $-$$ for additional services and data	https://www.ltcdatacooperative.org/Pages/default.aspx ; https://www.nia.nih.gov/research/dbsr/nia-data-linkage-program-linkage

Data Source	Description	Key Variables	Access (Public/Restricted) & Process	HIPAA-protected (yes/no)	Smallest Unit of Analysis (e.g., resident, facility)	Other Possible Units of Analysis	Linkable & At What Unit(s)	Benefits	Limitations	Use Case Examples	Cost (0=free, $=low, $$=moderate, $$$=high)	URL to Access Data
Long-Term Care Data Cooperative (LTCDC) - Unlinked	Nationwide database aggregating electronic health record (EHR) data from NHs. Data from multiple EHR vendors are harmonized into a single dataset.	Resident demographics, diagnoses, medication orders, lab results, care plans.	Restricted; requires LTCDC approval and a data use agreement.	Yes	Clinical measurement (e.g., medication administration, laboratory measure, vital sign)	Resident, Facility	Yes, resident-level	Rich clinical data in real-time, used for direct patient care; information absent from other datasets is available (e.g., laboratory values); more representative than single-vendor EHR data.	Data completeness varies by EHR vendor; potential missingness in fields; no-cost option requires data management expertise; use restricted to observational comparative effectiveness studies and clinical trials.	Analyzing care pathways, medication prescribing patterns, and resident outcomes.	0 for standard access; $-$$ for additional services and data	https://www.ltcdatacooperative.org/Pages/default.aspx
Long-Term Care Pharmacy (e.g., Omnicare, PharMerica) Dispensing Data	Proprietary pharmacy dispensing data collected by LTC pharmacies (e.g., Omnicare, PharMerica) for medications administered to NH residents. Structured similarly to Medicare Part D claims but includes medications regardless of payer.	Drug name, strength, formulation, route, quantity, days supplied, dispense date, refill status, resident ID (de-identified), facility and prescriber identifiers may be available upon request.	Restricted; no formal public access process. Requires individual negotiation and business/data use agreements with the pharmacy vendor.	Yes	Claim	Resident (and Prescriber and Facility if ID available)	Yes, resident-level (and potentially facility-level)	High granularity capture of dispensed medications; includes drugs dispensed during a Medicare Part A–covered SNF stay; valuable for real-world medication use and safety studies. It is also possible to link to CMS data.	No standardized public access; access depends on business partnerships; no data on clinical outcomes or diagnoses; scope limited to pharmacy's NH clients.	Evaluating changes in anticholinergic burden before and after hip fracture in SNF residents to identify deprescribing targets; Describing use of high-risk medications before and after hospitalization among NH residents using pharmacy and claims data.	$$/$$$	Not publicly available; access negotiated via vendor (e.g., https://www.omnicare.com, https://www.pharmerica.com)
Medicare Beneficiary Summary File (MBSF)	CMS file that includes demographic information, enrollment status, and derived indicators for all Medicare beneficiaries. These indicators summarize chronic conditions, disability status, dual eligibility, and mortality, based on claims data from the corresponding calendar year.	Beneficiary demographics (age, sex, race/ethnicity), enrollment status (Parts A, B, C, D), dual eligibility status, date of death, chronic condition and disability flags, geographic identifiers.	Restricted; requires CMS data use agreement and approval via ResDAC or as part of the NIA LINKAGE Program.	Yes	Resident	County, Facility, State, Region	Yes, resident-level, facility-level	Provides demographic, program enrollment, and comorbidity information for nearly all Medicare beneficiaries; useful for cohort building, risk adjustment, and identifying chronic or disabling conditions.	Contains limited clinical detail; some variables (e.g., race/ethnicity) may be misclassified; condition flags are based on claims-derived algorithms and not clinical records, so misclassification is possible.	Constructing longitudinal cohorts of NH residents, identifying dual-eligible beneficiaries, stratifying by comorbidity or disability status, assessing mortality and demographic patterns.	0/$	https://resdac.org/cms-data/files/mbsf-base
Medicare Compare (Nursing Home, Hospital, Hospice)	Publicly available healthcare facility performance metrics, including for nursing homes.	Star ratings, inspection results, staffing levels, quality measures.	Public; accessible via CMS website.	No	Facility	Market, State	Yes, facility-level	Facilitates comparisons of nursing home quality at the facility level.	Lacks resident-specific data, quality metrics reported by the facility, and information on valuable concepts like nursing home culture.	Analyzing trends in nursing home performance, regulatory compliance.	0	https://data.cms.gov/provider-data/?redirect=true
Medicare Hospice Claims	CMS administrative claims data capturing hospice care services provided to Medicare-covered residents, including information on timing, setting, and type of hospice care received.	Dates of hospice enrollment and discharge, type and level of hospice care, diagnosis codes, service location, provider identifiers, payments, resident demographics.	Restricted; requires CMS data use agreement and approval via ResDAC or as part of the NIA LINKAGE Program.	Yes	Claim	Resident, Facility, Hospice Provider	Yes, resident-level, facility-level	Enables analysis of hospice utilization, intensity, and timing among NH residents; can identify transitions to hospice and patterns of end-of-life care.	Limited to Medicare-covered hospice care; lacks detailed clinical context such as symptoms or preferences; services not billed to Medicare (e.g., Medicaid hospice) are not captured.	Examining hospice use before death in NHs, characterizing late vs. early hospice enrollment, evaluating variation in hospice transitions across facilities or regions.	$	https://resdac.org/cms-data/files/hospice-ffs
Medicare Part A / Medicare Provider Analysis and Review (MedPAR)	CMS administrative claims data capturing inpatient hospitalizations, skilled nursing facility (SNF) stays, hospice care, and other institutional services covered under Medicare Part A.	Admission and discharge dates, diagnosis and procedure codes (e.g., ICD-10), patient demographics, discharge destinations, provider identifiers, length of stay, charges, payment information.	Restricted; requires CMS data use agreement and approval via ResDAC or as part of the NIA LINKAGE Program.	Yes	Claim	Resident, Facility	Yes, resident-level, facility-level	Comprehensive, detailed claims data for Medicare-covered inpatient and institutional care; enables detailed analysis of hospitalization and SNF utilization patterns.	Limited to Medicare beneficiaries; subject to coding inaccuracies; incomplete for Medicare Advantage population; requires more expertise to analyze proficiently.	Examining hospitalizations among NH residents, analyzing post-acute care trajectories, evaluating associations between inpatient events and NH outcomes.	$$	https://resdac.org/cms-data/files/medpar

Data Source	Description	Key Variables	Access (Public/Restricted) & Process	HIPAA-protected (yes/no)	Smallest Unit of Analysis (e.g., resident, facility)	Other Possible Units of Analysis	Linkable & At What Unit(s)	Benefits	Limitations	Use Case Examples	Cost (0=free, $=low, $$=moderate, $$$=high)	URL to Access Data
Medicare Part B Carrier File	CMS administrative claims data capturing healthcare services provided by physicians, physician assistants, nurse practitioners, and other healthcare providers under Medicare Part B coverage.	Procedure codes (CPT/HCPCS), diagnosis codes (ICD-10), service dates, provider identifiers and specialties, place of service, charges, reimbursement amounts.	Restricted; requires CMS data use agreement and approval via ResDAC or as part of the NIA LINKAGE Program.	Yes	Line item	Resident, Facility	Yes, resident-level, facility-level	Detailed outpatient claims data capturing professional services, diagnostics, and treatments; essential for analyzing outpatient medical care patterns among NH residents.	Limited to Medicare fee-for-service beneficiaries, does not include data on beneficiaries enrolled in Medicare Advantage; data quality dependent on accuracy and consistency of provider coding.	Studying patterns of outpatient care, physician practice patterns, use of diagnostic procedures, and outpatient service utilization among NH residents; examining access and utilization of specialty medical care services	$$	https://resdac.org/cms-data/files/carrier-ffs
Medicare Part C (a.k.a. Medicare Advantage) Encounter (a.k.a. Carrier) Records	CMS data capturing healthcare encounters for beneficiaries enrolled in Medicare Advantage (Part C) plans. Contains encounter records analogous to Medicare Parts A and B claims, but submitted by Medicare Advantage plans.	Diagnosis and procedure codes (ICD-10, CPT/HCPCS), dates of service, provider identifiers, beneficiary demographics, place of service, encounter types, reimbursement information.	Restricted; requires CMS data use agreement and approval via ResDAC or as part of the NIA LINKAGE Program.	Yes	Line item	Resident, Facility	Yes, resident-level, facility-level	Provides detailed healthcare utilization data for Medicare Advantage beneficiaries; essential for analyzing healthcare patterns among a large and growing segment of NH residents enrolled in Medicare Advantage plans.	Data completeness and quality vary by Medicare Advantage plan; less standardized than traditional Medicare claims; encounter records may underestimate utilization due to variability in plan reporting practices.	Studying healthcare utilization, service patterns, and outcomes among NH residents enrolled in Medicare Advantage plans; examining differences in care received by NH residents under Medicare Advantage compared to traditional fee-for-service Medicare.	$$	https://resdac.org/cms-data/files/carrier-encounter
Medicare Part D Drug Claims	CMS administrative claims data for prescription drugs dispensed to Medicare beneficiaries enrolled in Part D. Includes detailed records of filled prescriptions.	National Drug Code (NDC), fill date, days supplied, quantity dispensed, drug name, cost, plan-paid amount, beneficiary-paid amount, pharmacy identifier, prescriber identifier.	Restricted; requires CMS data use agreement and approval via ResDAC or as part of the NIA LINKAGE Program.	Yes	Claim	Resident, Prescriber, Pharmacy, Facility	Yes, resident-level, facility-level	Provides detailed, standardized data on prescription drug utilization among Medicare beneficiaries; useful for examining medication access, effectiveness, and safety	Limited to Medicare Part D enrollees; does not capture medications during a Medicare Part A–covered SNF stay or hospitalization; may exclude over-the-counter drugs and drugs paid by other sources.	Estimating the effedts of medication regimens; describing high-risk medication use, polypharmacy, medication discontinuation, or trends in medication use among NH residents.	$$	https://resdac.org/cms-data/files/pde
Medicare Skilled Nursing Facility (SNF) Claims	CMS administrative claims data that capture services provided during Medicare Part A–covered post-acute SNF stays, including information on stay timing, reimbursement, and clinical diagnoses.	Dates of SNF admission and discharge, length of stay, diagnosis and procedure codes (e.g., International Classification of Diseases, 10th Revision [ICD-10]), Resource Utilization Group (RUG) codes, charges, payments, provider identifiers, resident demographics.	Restricted; requires CMS data use agreement and approval via ResDAC or as part of the NIA LINKAGE Program.	Yes	Claim	Resident, Facility, SNF Provider	Yes, resident-level, facility-level	Enables detailed tracking of Medicare-covered post-acute care utilization in SNFs, including timing, intensity, and payment structure; essential for studying care transitions and outcomes after hospitalization.	Limited to Medicare Part A–covered SNF stays; does not capture long-stay custodial NH care or services not billed to Medicare; clinical detail is limited to claims-derived codes.	Studying transitions from hospital to SNF, examining variation in SNF length of stay and payments, evaluating post-acute care patterns and outcomes among NH residents.	$	https://resdac.org/cms-data/files/snf-ffs
Minimum Data Set (MDS)	Standardized resident assessments measuring health status for all nursing home residents at admission, quarterly, and when there is a change in condition.	Demographics; functional status, cognitive performance, psychosocial functioning, diagnoses, medications, care needs.	Restricted; requires CMS approval and a data use agreement. Also possible to obtain directly from NH chain or as part of the NIA LINKAGE Program.	Yes	Assessment	Resident, Facility	Yes, resident-level	Rich longitudinal clinical data on nursing home residents collected from in all nursing homes that accept Medicaid and/or Medicare funding.	Limited medication details; some self-reported elements are incomplete or potentially biased; data collected quarterly.	Evaluating functional/cognitive status changes, quality of care monitoring.	$	https://resdac.org/cms-data/files/mds-30

Data Source	Description	Key Variables	Access (Public/Restricted) & Process	HIPAA-protected (yes/no)	Smallest Unit of Analysis (e.g., resident, facility)	Other Possible Units of Analysis	Linkable & At What Unit(s)	Benefits	Limitations	Use Case Examples	Cost (0=free, $=low, $$=moderate, $$$=high)	URL to Access Data
Online Survey Certification and Reporting / Certification and Survey Provider Enhanced Reporting (OSCAR/CASPER)	National dataset containing federally mandated NH survey and certification data. OSCAR was replaced by CASPER in 2012, improving data documentation.	Facility characteristics, staffing levels, deficiencies cited during inspections, penalties/enforcement actions, resident census information, facility ownership details.	Publicly available via CMS website. Historical OSCAR data may require special requests.	No	Survey	Facility, County, Market, State	Yes, facility-level	Provides comprehensive historical and current NH survey data; facilitates national comparison of NH quality, compliance, and regulatory outcomes.	Lacks resident-level clinical information; historical OSCAR data involved data overwriting, limiting longitudinal analyses; survey data may reflect subjective inspector judgments.	Assessing trends in NH regulatory compliance, examining staffing patterns, facility characteristics, and the relationship between deficiencies and quality outcomes.	0	
Veterans Affairs (VA) Community Living Center (CLC) Data (i.e., VA Corporate Data Warehouse, GECDAC residential history file, CLC dashboards, Support Service Center Capital Assets [VSSC] reports, VA Information Resource Center [VIReC]	Electronic health record data from the VA Corporate Data Warehouse (CDW), including detailed clinical and administrative information on residents who receive post-acute or long-term care in VA Community Living Centers (CLCs). Data are drawn from the broader VA health system and can be linked to inpatient, outpatient, pharmacy, laboratory, and administrative datasets.	Resident demographics, diagnoses, medication orders and administrations, vital signs, laboratory results, CLC stay timing and location, care plans, procedures, and other clinical and administrative data across VA settings.	Restricted; requires VA affiliation and Institutional Review Board (IRB) approval, along with approval from the VA Research and Development Committee. Access is typically limited to VA-employed or VA-affiliated investigators.	Yes	Clinical event (e.g., medication administration, lab result, vital sign)	Resident, Facility, Episode of Care	Yes, resident-level, facility-level	Enables comprehensive longitudinal tracking of Veterans across institutional, outpatient, and inpatient care settings; captures rich clinical detail for CLC residents; supports cohort construction and care trajectory analyses.	Generalizability may be limited due to Veteran population and VA-specific care delivery model; access restricted to VA-affiliated researchers; data may require significant technical expertise to extract and manage.	Building retrospective cohorts of Veterans receiving CLC care; analyzing longitudinal care patterns across VA settings; evaluating medication use and effects, care transitions, and outcomes among residents.	0 (for VA-affiliated researchers)	https://www.virec.research.va.gov/

Funded by U24AG087939, NIA