Using Natural Language Processing to Identify Home Health Care Patients at Risk for Diagnosis of Alzheimer's Disease and Related Dementias

Miriam Ryvicker, PhD[a], **Yolanda Barrón, MS**[a], **Jiyoun Song, PhD, AGACNP-BC, RN**[b], **Maryam Zolnoori, PhD**[c], **Shivani Shah, MPH**[a], **Julia Burgdorf, PhD**[a], **James M. Noble, MD, MS, CPH, FAAN**[d], **Maxim Topaz, PhD, RN, MA, FAAN, FIAHSI**[a,d]

[a.]Center for Home Care Policy & Research at VNS Health

[b.]University of Pennsylvania School of Nursing

[c.]Columbia University School of Nursing

[d.]Columbia University Medical Center

Abstract

Objectives: This study aimed to: (1) validate a natural language processing (NLP) system developed for the home health care setting to identify signs and symptoms of Alzheimer's Disease and related dementias (ADRD) documented in clinicians' free-text notes; (2) determine whether signs and symptoms detected via NLP help to identify patients at risk of a new ADRD diagnosis within four years after admission.

Methods: This study applied NLP to a longitudinal dataset including medical record and Medicare claims data for 56,652 home health care patients and Cox proportional hazard models to the subset of 24,874 patients admitted without an ADRD diagnosis.

Results: Selected ADRD signs and symptoms were associated with increased risk of a new ADRD diagnosis during follow-up, including: motor issues; hoarding/cluttering; uncooperative behavior; delusions or hallucinations; mention of ADRD disease names; and caregiver stress.

Conclusions: NLP can help to identify patients in need of ADRD-related evaluation and support services.

Keywords

Alzheimer's Disease; dementia; home health care; natural language processing

---

Introduction

Alzheimer's disease and related dementias (ADRD) affect roughly 5 million people and 11% of older adults in the United States (U.S.) (Alzheimer's Association, 2022; Matthews et al., 2019). Pathways to ADRD diagnosis and treatment vary, and delayed diagnosis is often a barrier to accessing treatment for individuals with ADRD and support services for their families (Hinton, Franz, & Friend, 2004; Lin et al., 2021; Speechly, Bridges-Webb, & Passmore, 2008). Older adults with ADRD tend to be clinically complex; roughly half have three or more chronic conditions (Lin, Fillit, Cohen, & Neumann, 2013). Given their clinical complexity, many patients with ADRD or who may be on a path toward ADRD diagnosis may be admitted for home health care services to meet an array of skilled care needs (Ryvicker et al., 2022; Topaz, Adams, Wilson, Woo, & Ryvicker, 2020).

Home health care is a setting where skilled clinicians, including nurses, physical and occupational therapists, and social workers, provide episodic healthcare services to patients in their homes, with additional support for activities of daily living provided by home health aides to eligible patients. In the U.S., 12,000 home care agencies serve more than 6 million patients annually (Romagnoli, Handler, & Hochheiser, 2013). Home care clinicians are uniquely situated to detect symptoms and provide tailored interventions for ADRD patients or referrals for patients needing comprehensive cognitive evaluation. However, little is known about home care clinician practices in documenting their patients' cognitive status. Information in the electronic health record, particularly clinicians' free-text notes, can provide insights into home care clinicians' knowledge of patients' prior ADRD diagnoses, observations of cognitive symptoms and caregiver needs, and related interventions.

Natural language processing (NLP) is a body of methods for systematically analyzing free-text content. It can be used to explore untapped, yet critical patient data found in clinical notes. NLP is being used increasingly in healthcare, with up to 80% of information stored as free text (e.g. progress/follow-up notes, admission/discharge summaries, radiology reports) (Ford, Carroll, Smith, Scott, & Cassell, 2016; Hassanpour & Langlotz, 2016; Meystre, Savova, Kipper-Schuler, & Hurdle, 2008). ADRD involves a range of behavioral and emotional signs and symptoms, complex family dynamics, and a long trajectory of changes over time (Fisher, Schwartz, Greenspan, & Heinrich, 2016; Gitlin et al., 2018; Gitlin, Kales, Marx, Stanislawski, & Lyketsos, 2017; Sloane et al., 2017; Sperling et al., 2011; Stansfeld et al., 2018; Werner, Friedland, & Inzelberg, 2015). Home care clinicians' notes are therefore potentially rich in narrative, with vocabularies including both explicit and implicit references to ADRD. Moreover, observations of patients in their home environment may highlight ADRD signs, symptoms, and unmet needs which may be difficult to observe in other healthcare settings, such as a hospital or outpatient medical office, outside of the context of daily routines. Although NLP could be useful in identifying clinicians' observations of signs and symptoms, it has been underutilized in ADRD research within the home care setting.

In previous pilot work, our study team developed an NLP vocabulary to identify neuropsychiatric symptoms of dementia in a retrospective sample of patients from a large, non-profit home care organization (Topaz et al., 2020). The objectives of the current study are to: (1) further refine and validate this NLP system in a retrospective cohort of

56,652 home care patients with and without known ADRD diagnoses upon home care admission; and (2) using data from a subset of this cohort who did not have an established ADRD diagnosis upon home care admission (N=24,874), determine whether NLP-derived indicators of ADRD constructs during the home care episode help to predict a subsequent ADRD diagnosis during a four-year period following admission, thus helping to identify individuals in need of ADRD-related diagnostic and support services.

Methods

An observational retrospective cohort study with secondary data analysis was conducted using structured data (i.e., demographic and clinical assessment data) and unstructured data (i.e., clinical notes). The cohort consisted of 56,652 home care patients served by VNS Health, a large non-profit home care provider in the U.S. The study was approved by the Institutional Review Board of VNS Health (Protocol # 1585549-1).

Data sources.

The data included the federally mandated Outcome and Assessment Information Set (OASIS) clinical assessment administered at the start of a home care episode, linked with Medicare claims data for a 6-month look-back period and 4 years of follow-up. The home care episode refers to a period of up to 60 days during which home care services are provided. Claims data spanned the calendar years of 2010–2016, allowing uniform observation periods. The four-year follow-up period was selected to allow for the maximum length of time available for the study cohort at the time of data acquisition, acknowledging the long disease trajectory of ADRD.

OASIS is a standard assessment tool mandated by the Centers for Medicare & Medicaid Services (CMS) with roughly 100 items on clinical status and service needs during a home care episode. OASIS includes data on patient demographics, living arrangements, informal supports, co-morbidities, symptom severity, risk factors, prognosis, therapies, medication/equipment management, pain, wounds, neuro-cognitive/behavioral status, and physical function (Kinatukara, Rosati, & Huang, 2005; Tullai-McGuinness, Madigan, & Fortinsky, 2009). OASIS data and language information were extracted from the VNS Health electronic health record and linked with Medicare enrollment and claims data acquired through a Data Use Agreement from CMS. These data included demographics, insurance enrollment information, chronic condition diagnoses, and claims for all services provided under Fee-for-Service Medicare Parts A and B in inpatient (e.g. hospitals, skilled nursing facilities), outpatient, office, home care and hospice settings. All claims files contained service dates, diagnoses, and procedure codes.

Approximately 1.8 million free-text clinical notes were extracted from the home care electronic health record for the study cohort. The narrative notes were completed by home care clinicians, including registered nurses, social workers, physical and occupational therapists, and clinical team managers. Clinical notes included: (1) visit notes that describe the patient's conditions and care provided during home care visits (n=788,564); and (2) care coordination notes describing communications between clinicians and other administrative activities related to patient care (n=973,470). Patients had an average of 9.4 total visit notes

per episode of care (standard deviation=6.8; range: 1–31), and an average of 13.1 care coordination notes (standard deviation= 7.3; range: 1–31).

Inclusion and exclusion criteria.

Patients were included in the full study cohort (N=56,652) if they were admitted to home care at VNS Health from July 1, 2010 through December 31, 2012 with any diagnosis and had Medicare Fee-for-Service coverage for the entire 6-month look-back period and 4-year follow-up period. Patients were excluded from the sample if they were enrolled in Medicare Advantage at any point during the lookback or follow-up period, since the dataset did not include Medicare Advantage claims. The full cohort (N=56,652) was used to develop and refine the NLP vocabulary. To examine associations between NLP-derived indicators related to ADRD, we restricted the sample to patients who did not have an established ADRD diagnosis prior to home care admission and either: (1) were diagnosed with ADRD during the 4-year follow-up period; or (2) were alive at the end of the 4-year follow-up and were not diagnosed with ADRD during that timeframe. We identified ADRD diagnoses using the list of ICD codes included in the ADRD indicator provided by CMS in the Chronic Condition Warehouse, including the relevant ICD-9 and ICD-10 codes for their respective timeframes during the study period (Chronic Condition Data Warehouse, 2018a, 2018b). We excluded patients who died during the 4-year follow-up period without a new ADRD diagnosis prior to death, since the outcome could not have been observed in these patients. The above criteria yielded an analytic sample of N=24,874; all results reported in this paper reflect this analytic sample.

NLP algorithm development.

Our team developed an NLP algorithm to extract ADRD-related concepts from the free-text clinical notes. The NLP development and validation were carried out in a 5-step process, depicted in Figure 1. First, based on a comprehensive literature review and discussions with the study's clinical experts, our interdisciplinary team identified concepts related to three information domains that potentially indicate risk for ADRD: ADRD disease names; signs/symptoms; and interventions. We then specified 28 subcategories (e.g., 'mentions of ADRD,' 'memory loss,' 'caregiver stress,' 'addressing unmet needs') under the three domains. Second, we created a preliminary list of terms and expressions for each subcategory based on standard terminology – including the Unified Medical Language System (UMLS) (Pradeep, Sunder, Bendale, Mantri, & Dande, 2013) and Systematized Nomenclature of Medicine (SNOMED) (Stearns, Price, Spackman, & Wang, 2001). Third, we expanded the synonyms for the preliminary list of terms, using a specific type of language model called word embedding, or Word2Vec, which detects synonyms for terms of interest within specific domains by learning word associations based on a large corpus of text. For this task, we used *NimbleMiner* (Topaz, Murga, Bar-Bachar, McDonald, & Bowles, 2019), an open-source NLP application in RStudio (The R Foundation, 2023).

Fourth, we implemented an interactive vocabulary explorer to identify synonymous expressions, also using NimbleMiner. This process returns suggested synonyms, including potential misspellings; for example, a query for synonyms of the term "poor sleep" returns terms including "altered sleep" and "chronic insomnia" and potential misspellings including

"inosmania" or "sleeplessensse." The process also accounts for potential negations; for example, words such as "no" and "deny" were identified in NimbleMiner to indicate whether words or expressions might have been negated our ruled-out, such as "reports *no* insomnia" or "*denies* problems with sleep." Two team members [JS and MT] reviewed the results generated by the interactive vocabulary explorer user interface and decided whether to accept or reject them. The process was repeated until no new synonyms were identified.

Finally, we evaluated the performance of this NLP algorithm for automatic identification of ADRD risk factors from clinical notes. For each category of ADRD-related concepts, 30 clinical notes in which the algorithms identified subcategories were extracted to determine whether the algorithms correctly identified the information. The NLP system showed excellent performance in identifying the ADRD-related concepts, with an average precision of 0.945, ranging from 0.76 to 1. Further detail about the development and evaluation of the NLP system, concept definitions, and system performance is described in the Appendix.

Association analysis of NLP-derived risk factors and ADRD.

Having applied the NLP algorithm to all the clinical notes, a binary variable was generated to indicate whether each of the ADRD-related constructs was documented for each patient. We then used Cox proportional hazards regression models to estimate associations of the ADRD-related constructs with the time to ADRD diagnosis during the 4-year follow-up period among patients without an established ADRD diagnosis prior to home care admission. We tested the proportionality assumption for each independent variable in the model – namely, the assumption that the effect of the variable remained constant throughout the 4-year follow-up period. For each independent variable where the proportionality assumption did not hold, we estimated Cox models with time-varying coefficients. To do so, we created interaction terms between the variables where the proportionality assumption did not hold, and binary variables indicating the follow-up year indexed on the home care admission (from 1 to 4 years) and re-estimated the models. These models estimated different hazard ratios for each of the 4 years of follow-up. We checked that the proportionality assumption held after the time-varying coefficient estimation. For each variable where the proportionality assumption held, the model estimated a unique hazard ratio for the entire 4-year follow-up period.

The model was case-mix adjusted for age, sex, race, type of facility from which the patient was referred to home care, and several clinical and functional items on the structured OASIS assessment upon home care admission. These items included: falls history; recent mental decline; overall prognosis; pain; dependencies in activities of daily living (ADLs); hearing impairment; comprehension of verbal content; speech impairment; urinary and bowel incontinence; surgical wounds; structured assessment of cognitive function and confusion; ability to manage medications; supervision and safety assistance needs; and plan of care orders for diabetic foot care and pain management. We used RStudio version 3.6 for this analysis.

Results

Patient characteristics.

The analytic sample included 24,874 patients who were either: (i) without an ADRD diagnosis and alive at the end of the 4-year follow-up period (64%); or (ii) diagnosed with ADRD during the follow-up period (36%). Selected characteristics indicated in the OASIS assessment at home care admission are shown in Table 1. Patients who were subsequently diagnosed with ADRD tended to be older (mean age of 82 vs 73 years old) and were more likely to be non-Hispanic White (68% vs 65%). They were more likely to have documentation of prior impaired decision-making (5.4% vs 2.6%), memory loss (3.5% vs 0.8%) and urinary incontinence (18.6% vs 9.7%). They were also identified more often as being at risk for hospitalization due to recent mental decline (12.9% vs 7.5%), frailty (30.2% vs 22.8%), or a history of falls (20.8% vs 11.3%). All items were statistically significant at the p-value level <0.001, except for sex (p-value = 0.048).

Timing of ADRD diagnosis.

The distribution of the timing of ADRD diagnosis during the 4-year follow-up period is shown in Table 2. Of the 8,901 patients diagnosed with ADRD during follow-up, 40.0% were diagnosed in Year 1, 25.3% in Year 2, 20.6% in Year 3, and 14.1% in Year 4. This distribution provided sufficient variation in the timing of the outcome for the Cox proportional hazards model.

NLP findings of ADRD-related constructs.

The presence of ADRD-related constructs detected in the clinical notes via NLP is shown in Table 3 for the overall sample and with comparisons by ADRD diagnosis during the follow-up period, along with examples of each construct. Overall, ADRD constructs ranged from very rare to quite common. For example, only 0.2% of the overall sample were noted to experience wandering behavior, whereas 10.6% were noted as having memory loss, 19.6% having signs of anxiety, and 57.7% receiving one or more home safety intervention(s).

By and large, the ADRD constructs identified in the notes appeared more frequently among patients who were diagnosed with ADRD during the follow-up period. Differences in the presence of these constructs were statistically significant at p<0.001 for all variables except: wandering (p=0.002); functional issues (p=0.137); and interventions to address unmet needs (p=0.954). Compared to those not diagnosed during follow-up, patients diagnosed with ADRD were more likely to have clinical notes mentioning ADRD (5.1% vs 1.9%), memory loss (17.6% vs 6.7%), agitation (4.3% vs 2.0%), and confusion (7.5% vs 2.6%). They were also more likely to have notes mentioning uncooperative behavior (10.1% vs 5.6%), motor issues (14.7% vs 9.0%), and delusions or hallucinations (1.2% vs 0.4%).

The groups with and without a new ADRD diagnosis during the follow-up period also differed in their overall patterns in the potentially ADRD-related symptoms most commonly documented. The five most commonly documented symptoms in the group with a new diagnosis were: anxiety (22.4%); memory loss (17.6%); motor issues (14.6%); depression or apathy (14.0%); and uncooperative behavior (10.1%). While anxiety was also the most

commonly documented symptom in the group without a new ADRD diagnosis (18.1%), the other most common symptoms differed: depression or apathy (10.1%); motor issues (9.0%); memory loss (6.7%); and the fifth most common tied between uncooperative behavior and physical or verbal aggression (both at 5.6%).

Multivariate Cox proportional hazards model with time-varying coefficients.

Findings from the Cox proportional hazards model show that several NLP-derived indicators of ADRD-related signs and symptoms were significantly associated with increased risk, as indicated by the hazard ratio, of receiving a new ADRD diagnosis with 4 years of home care admission (Table 4). Mentions of motor issues, hoarding and clutter, uncooperative behavior, and delusions or hallucinations were all associated with an increased risk of being newly diagnosed with ADRD across the entire 4-year follow-up. The proportionality assumption held for these variables; thus, these items did not have time-varying effects.

Among the time-varying effects, several NLP-derived indicators had a significant effect during some follow-up years and not others (Figures 2 and 3). A clinician's mention of ADRD in the notes was associated with an increased risk of receiving an ADRD diagnosis in years 1 and 2, with the effect no longer significant in years 3 and 4. Mentions of caregiver stress were associated with a 1.695 increase in the risk of ADRD diagnosis within 1 year of home care admission; however, this effect did not remain significant in years 2 through 4. The association of memory loss with increased risk of ADRD diagnosis remains significant across all 4 years, with the strongest effects in years 2 and 3 (Figure 4).

Discussion

This study is the first, to our knowledge, to develop an NLP system within the home care setting that detects ADRD-related constructs in home care clinicians' free-texts notes and examine associations of these constructs with subsequent ADRD diagnosis. Other studies have developed NLP systems to identify dementia-related constructs within primary care and other outpatient settings, hospitals, and healthcare systems (Hane, Nori, Crown, Sanghavi, & Bleicher, 2020; Maclagan et al., 2023; Noori et al., 2022; Oh et al., 2023). However, findings from these studies are not directly comparable to the findings presented here, given differences in setting, design and methodologies. The NLP system developed in the current study detected a substantial volume of free-text documentation of home care clinicians' observations pertinent to ADRD, including home safety issues, caregiver stress, and other phenomena that may not be easily observed in office-based or inpatient settings.

Mentions of terms alluding to ADRD disease names were relatively rare in the clinical notes. This is not surprising given that the home care clinician's role does not include making new diagnoses. Nevertheless, a clinician may document observations or reports from family members who discuss the patient's cognitive status in terms of dementia, regardless of whether the person has received or is expected to receive a formal ADRD diagnosis that would appear on a medical claim. It is possible that, in some instances, family members will express that the patient has some form of dementia, without accessing services that would lead to a formal diagnosis. This is consistent with prior research indicating that normalization of dementia symptoms and lack of perceived need are barriers to seeking

help for a dementia diagnosis (Parker, Barlow, Hoe, & Aitken, 2020). Home care clinicians' free-text documentation may in part reflect family members' understanding of the person's cognitive status, even without a formal diagnosis on record.

The ADRD symptom domain contained the largest volume of constructs identified by the NLP system. Across all three domains, home safety interventions comprised the single most frequently observed construct across groups with and without an incident ADRD diagnosis (63.3% and 54.6%, respectively). The frequent mention of home safety interventions is not surprising, given that safety assessments are a key component of home care services (Romagnoli et al., 2013). In the ADRD symptom domain, the group of patients with an incident ADRD diagnosis had a somewhat different set of the top five most commonly identified symptoms than those without an incident diagnosis. Although some of the commonly identified symptoms are not unique to ADRD – such as anxiety, depression/apathy, and motor issues – the higher concentration of these symptoms in the group with an incident diagnosis suggests that the NLP system can detect signals of heightened risk for ADRD diagnosis.

In case-mix adjusted models, some NLP-derived indicators had statistically significant associations with new ADRD diagnosis, with constant relative hazards across all four years of follow-up. Specifically, motor issues, hoarding and cluttering, uncooperative behavior, and delusions or hallucinations met the proportionality assumption. Although we can speculate as to why these items were consistently associated with heightened risk of an ADRD diagnosis regardless of the point of follow-up, we do so with caution. It is possible that these symptoms are more commonly observed and documented for individuals in more advanced stages of the disease, and/or these symptoms are more specific to ADRD than some of the other symptoms of interest. We also note that while the proportionality assumption did not hold for the documentation of memory loss, this item was significantly associated with ADRD diagnosis across all 4 years of follow-up, but with slightly different relative hazards. This is notable because memory loss is likely a robust signal that a neuropsychiatric evaluation is needed (Fisher et al., 2016).

Other NLP-derived indicators with time-varying effects included the mention of at least one ADRD disease name and caregiver stress. The mention of an ADRD disease name was associated with a 27% increase in the risk of being diagnosed with ADRD within the first year after home care admission, and a 40% increase in year 2. Additionally, the mention of caregiver stress was associated with a 70% increase in the risk of a new ADRD diagnosis within the first year after admission. These findings suggest that, even though it is not within the home care clinician's scope of work to make a new diagnosis, a clinician may be well positioned to detect when a patient is on a path toward potential diagnosis, as well as when a caregiver needs additional support related to the patient's cognitive decline. Although it is unclear why these effects were no longer significant in later years, it is possible that the clinician's mention of ADRD terms and/or caregiver stress signals a period of heightened concern for the patient's cognitive status that precedes the pursuit of ADRD diagnostic services. The decline in effect significance in later years may be due to progressively smaller cell sizes in the subgroups diagnosed with ADRD in Years 2, 3, and 4. It should also be

noted that the time-varying effects for some other variables, such as agitation, had irregular patterns across follow-up years, which may also have been influenced by small cell sizes.

The signals of heightened ADRD risk detected by the NLP system have potential utility in identifying individuals at different stages on the ADRD disease trajectory who may be in need of evaluation for ADRD, and/or may have dementia-related needs that can be addressed in the home during the months or years prior to a formal diagnosis. This is important given that many individuals experience delayed diagnosis (Hinton et al., 2004; Lin et al., 2021) but have existing needs related to cognitive decline which should be addressed to ensure safety, improve caregiver support, and facilitate advance care planning, without needing a formal medical diagnosis of ADRD as a trigger (Amjad et al., 2018). Moreover, since diagnostic practices vary, and access to gold standard diagnostic tools may be disparate (Cooper, Tandy, Balamurali, & Livingston, 2010), having mechanisms to identify and address these needs is essential to improving support for individuals and families more likely to be underserved in evaluation and diagnostic services (Amjad et al., 2018; Bradford, Kunik, Schulz, Williams, & Singh, 2009; Lin et al., 2021).

Some study limitations are noted. First, the study was limited to a single large home care provider; this may limit generalizability to other providers and geographies. The focus on the one provider supported a larger study, in which the diverse VNS Health population was advantageous for enhancing generalizability to understudied groups. A subsequent study can reproduce these methods with data from additional agencies in different regions to demonstrate reproducibility. Second, the retrospective dataset was developed for a previous study completed in 2018, using the most recent data available for the intended multi-year design at the time of data acquisition. Consequently, the structured assessment data was derived from an older version of the OASIS instrument (Centers for Medicare & Medicaid Services, 2022). Some of the final covariates selected from the structured data might have been different with more recent data. Third, patients were excluded if they enrolled in Medicare Advantage during the study period, since managed care claims were not available in the data use agreement. Fourth, in the development of our NLP algorithm, we utilized the word2vec methodology to identify synonyms. While this technique is efficient, transformer-based language models may offer more accurate results. Our future research will leverage these advanced methods to enhance our NLP algorithm. Finally, some of the time-varying coefficients generated by the Cox proportional hazards regression proved difficult to interpret, with differences in effects across follow-up years that may have been influenced by small cell sizes. Nevertheless, we found this to be the most appropriate method given the long follow-up period, and the findings on time-varying effects may be useful to researchers interested in testing a similar NLP model with a shorter follow-up period.

Conclusion

The findings of this study suggest that the free-text clinical notes documented by clinicians in the home care setting are a rich source of information on potential signs and symptoms of ADRD that can be detected using NLP methodology. Moreover, specific NLP-derived indicators of ADRD signs and symptoms can help to identify patients at heightened risk for

an eventual ADRD diagnosis, who therefore may be referred to dementia-related diagnostic services, and to address unmet needs for these patients and their caregivers. Home care clinicians are uniquely positioned to observe signs and symptoms that are evident in the home that may not easily be observed in other healthcare settings, and potentially to identify needs for specialist referrals (e.g. neurology, psychiatry), home safety interventions, and caregiver support services early in the ADRD disease trajectory. The NLP methodology could be used to efficiently and systematically identify individuals in need of diagnostic services, thus potentially mitigating the problem of delayed diagnosis and reducing its burdensome consequences (Amjad et al., 2018; Lin et al., 2021).

Further work is needed to develop and test strategies of integrating the NLP algorithm into home care clinical practice in real time, such that free-text documentation of ADRD constructs may be used to flag at-risk individuals for further care coordination services and referral for cognitive evaluation. These strategies would involve a coordinated effort between clinical and information technology teams within home care provider organizations to integrate the NLP model into the electronic health record, including an alert mechanism that informs the home care clinician that the patient may need cognitive evaluation and/or the family may need caregiver support services. In addition to these technological developments, home care organizations could increase the use of social work visits for flagged patients to facilitate referrals to community-based neurology/psychiatry specialists and caregiver support services (e.g. support groups, respite care). These strategies could include an adaptation of a home-based care coordination program designed for individuals with dementia (Samus et al., 2014). This approach holds promise for improving care delivery for individuals at different stages along the ADRD trajectory, facilitating access to caregiver support services and improving quality of life in this vulnerable population.

Acknowledgments:

Appendix: NLP Development Process

The overall process of developing NLP algorithms is depicted in Figure 1 in the main body of the manuscript. The following describes each step in this process in further detail.

(1) Identifying ADRD-related concepts that might appear in clinical notes

Based on a comprehensive literature review and several rounds of discussions with study experts, our interdisciplinary team (with expertise in home care nursing, medicine, social work, public health and bioinformatics) identified concepts related to three broad information domains that can be potentially indicative of risk for ADRD, including: ADRD disease names, signs/symptoms, and interventions. Furthermore, we specified 28 subcategories under the three information domains; definitions for each subcategory are shown in Table A.

(2) Creating a preliminary list of terms and expressions for each subcategory

As part of developing the preliminary lexicon for each subcategory, we used large standardized vocabularies of medical terms, including: Unified Medical Language System (UMLS) (Pradeep, Gaur, Prashant, Manisha, & Atreya, 2013), Systemized Nomenclature of Medical Terms (SNOMED) (Stearns, Price, Spackman, & Wang, 2001), International Classification of Diseases version 10 (ICD-10) (Quan et al., 2008), and International Classification for Nursing Practice (ICNP) (Wake & Coenen, 1998). Based on these standardized vocabularies, an initial list of terms was developed by four members of the team [JS, VA, PW, and MT], two of whom are home care nurses and additional two members who are PhD-prepared researchers in nursing informatics. Then the terms were reviewed and refined by all team members. For example, for the sign/symptom of "wandering," we identified 12 UMLS synonym expressions including "aimless walking" and "physical wandering," among others.

(3) Language model creation: word embedding model (Word2Vec)

Language models are statistical representations of a certain body of text and vectorized (numeric) representations of texts in specific domains that allow multiple NLP tasks, such as synonym detection and lexicon creation. A specific type of language model called word embedding (Word2Vec) was used in this study because it was found effective for lexicon discovery in previous studies (T. A. Koleck et al., 2021; Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, & Jeffrey Dean, 2013; T. Mikolov, I. Sutskever, K. Chen, G. Corrado, & J. Dean, 2013; Song et al., 2022; Topaz, Woo, Ryvicker, Zolnoori, & Cato, 2020a). When training Word2Vec, computers learn word associations from a large corpus of text. Once trained, Word2Vec models can be used to detect synonyms for terms of interest within specific domains. In this study, NimbleMiner (Topaz, Murga, Bar-Bachar, McDonald, & Bowles, 2019), an open-source publicly available NLP application in RStudio (Foundation of Statistical Computing, Vienna) was used to identify synonyms for subcategories present in a large body of home care clinical notes available for the study period. Earlier studies evaluated and validated the accuracy of NimbleMiner extracting symptom and important patient information from narrative clinical notes (Theresa A. Koleck et al., 2021; Topaz et al., 2016; Maxim Topaz et al., 2019; Topaz, Woo, Ryvicker, Zolnoori, & Cato, 2020b). The system can be downloaded from http://github.com/mtopaz/NimbleMiner under General Public License v2.0.

(4) Implementing an interactive vocabulary explorer to identify synonymous expressions

Based on the language model built in the previous step, we used NimbleMiner's "Interactive rapid vocabulary explorer" module that suggested synonyms based on a target words presented by the user. For example, when the system is queried for synonyms of the term "poor sleep," it returns terms such as "altered sleep" and "chronic insomnia" and potential misspellings such as "inosmania" or "sleeplessensse." Two reviewers [JS and MT] selected

the words via the interactive vocabulary explorer user interface and decided whether to accept or reject them. The process was repeated until no new synonyms were identified. Additionally, words such as "no" and "deny" were identified in the NimbleMiner application to indicate whether words or expressions might have been negated our ruled-out, such as "reports no insomnia" or "denies problems with sleep."

(5)   Evaluating NLP algorithm performance

Using NimbleMiner's "data labeling" module, NLP algorithm was applied on all clinical notes available for the study. To evaluate NLP algorithm's accuracy in identification of ADRD-related concepts, 30 random clinical notes for each information domain subcategory were selected for expert accuracy review. Each clinical note was evaluated by two team members [JS and MT] to determine if the algorithms identified the information regarding subcategories correctly or if there were errors in the identified instance. These results were used to calculate precision (akin to positive predictive value, calculated as the number of true positives out of the true positives plus false negatives]) for each category and to calculate an overall precision value. Recall (the number of true positives out of the actual number of positives) is another metric that is sometimes used to evaluate NLP accuracy. To calculate recall, a randomly extracted set of notes annotated by domain experts is needed. However, we could not calculate recall for this study since several subcategories of ADRD-related information domains were only very infrequently documented; for example, caregiver stress was only documented 44 times in all 1.8 million notes (results are further described in Table A). Hence, calculating recall would require pulling and reviewing an improbably large sample of random notes to ensure an inclusion of several mentions of caregiver stress. Based on our previous results of high recall in similar studies conducted with the same NLP approach (Chae, Song, Ojo, & Topaz, 2021; T. A. Koleck et al., 2021; Topaz et al., 2020), we felt confident that calculating precision is a good accuracy metric for this study.

The average precision for ADRD-related categories was 0.945 (range 0.76 – 1), indicating a high level of precision. The highest precision for subcategories included "ADRD disease names," "Memory loss," "Specific medications" and "Home safety interventions" (precision= 1); while the lowest precision was for "Functional issues" (0.76). See Table A for detail.

Table A.

ADRD construct definitions, frequency, and NLP performance (N=1,762,034 notes)

Domain / Construct	Definition	Frequency of notes detected by NLP algorithm	Frequency of notes confirmed by human review	Precision
Domain: ADRD disease names				
Mentions of ADRD	Presence of language indicative of ADRD disease names, including "Alzheimer's disease", "Dementia", "senile", etc.	673	30/30	1
Domain: Symptoms				

Domain / Construct	Definition	Frequency of notes detected by NLP algorithm	Frequency of notes confirmed by human review	Precision
Memory loss	Presence of language indicative of long term or short term memory problems, such as regularly forgetting recent events, names and faces, recognizing faces or family members	1717	30/30	1
Unspecified cognitive deficits	Presence of language indicative of general cognitive deficits, such as cognitive impairment	275	30/30	1
Communication problems	Any difficulty in communication, such as impaired speech, spouse answers all the questions	252	29/30	0.97
Motor issues (motor hyperactivity, motor slowness and other motor issues)	Presence of language indicative of motor issues, such as slowness, hyperactivity (e.g., tremor or shaking), falls, etc.	839	27/30	0.9
Physically/verbally aggressive	Presence of language indicative of verbal or physical aggression	810	29/30	0.97
Poor sleep	Presence of language indicative of sleep problems (e.g., insomnia)	322	29/30	0.97
Wandering	Presence of language indicative of wandering	34	26/30	0.86
Disorientation	Presence of language indicative of disorientation to place or person	101	27/30	0.9
Hoarding and clutter	Presence of language indicative of clutter or hoarding in the patient's apartment	316	29/30	0.97
Uncooperative/ refusal/non-compliance	Presence of language indicative of noncompliance (e.g., not taking medications, refusal of care, uncooperative)	1152	30/30	1
Depression/apathy	Presence of language indicative of depressive symptoms or depression diagnosis	2178	30/30	1
Anxiety	Presence of language indicative of anxious and restless behaviors	3215	30/30	1
Delusions/ hallucinations	Presence of language indicative of visual or hearing hallucinations or delusions	102	28/30	0.93
Agitation	Presence of language indicative of agitation	407	29/30	0.97
Confusion	Presence of language indicative of confusion and related symptoms	682	28/30	0.93
Bladder/bowel incontinence	Presence of language indicative of bladder/ bowel incontinence	555	29/30	0.97
Weight loss	Presence of language indicative of loss of weight (e.g., cachexia, anorexia)	449	29/30	0.97
Poor/increased appetite	Presence of language indicative of appetite issues	325	28/30	0.93
Trouble eating or swallowing	Presence of language indicative of problems with eating or swallowing	121	30/30	1
Executive function problems	Presence of language indicative of variety of issues, including problems with: verbal reasoning, problem-solving, planning, the ability to maintain sustained attention, resistance to interference, multitasking, cognitive flexibility, etc.	142	24/30	0.8

Domain / Construct	Definition	Frequency of notes detected by NLP algorithm	Frequency of notes confirmed by human review	Precision
Functional issues (ADL/IADL function issue)	Presence of language indicative of functional issues such as: problems with choosing clothing, needing clues and reminders in the bathroom and kitchen, leaving the stove on, problems with bathroom hygiene routines, etc.	103	23/30	0.76
Hearing loss	Presence of language indicative of mentions of hearing loss, hearing aids use	117	29/30	0.97
Caregiver stress	Presence of language indicative of increased needs of caregiver, such as increased caregiver strain and burden	44	26/30	0.86
Domain: Interventions				
Addressing unmet service needs	Presence of language indicative of addressing unmet service needs, e.g. referral to social work or psychiatry, teaching family about medication, increased home health aide hours	37	27/30	0.9
Referrals for end-of-life care	Presence of language indicative of referral for hospice care, advanced directives, end of life care, DNR	474	28/30	0.93
ADRD-specific medications	Presence of language indicative of specific medications that are often given to patients with ADRD-like symptoms, such as Donepezil, Galantamine, Memantine, Rivastigmine, etc.	414	30/30	1
Home safety interventions	Presence of language indicative of presence of assistive device, referral to superintendent in the building, safety hazards at home, hand rails, grab bars in the shower, door locks/alarms, GPS/tracking, safe return bracelet, etc.	10713	30/30	1
Overall precision				**0.945**

References

Alzheimer's Association. (2022). 2022 Alzheimer's Disease Facts and Figures. Retrieved from https://www.alz.org/media/documents/alzheimers-facts-and-figures.pdf

Amjad H, Roth DL, Sheehan OC, Lyketsos CG, Wolff JL, & Samus QM (2018). Underdiagnosis of Dementia: an Observational Study of Patterns in Diagnosis and Awareness in US Older Adults. J Gen Intern Med, 33(7), 1131–1138. doi:10.1007/s11606-018-4377-y [PubMed: 29508259]

Bradford A, Kunik ME, Schulz P, Williams SP, & Singh H (2009). Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors. Alzheimer Dis Assoc Disord, 23(4), 306–314. doi:10.1097/WAD.0b013e3181a6bebc [PubMed: 19568149]

Centers for Medicare & Medicaid Services. (2022). OASIS Data Sets. Retrieved from https://www.cambridge.org/core/journals/internationalpsychogeriatrics/article/abs/persistent-barriers-and-facilitators-to-seeking-help-for-a-dementia-diagnosis-a-systematic-review-of-30-years-of-the-perspectives-of-carers-and-people-with-dementia/F5A64C217D4484DB24CF14D24D1B81C5

Chronic Condition Data Warehouse. (2018a). CMS Chronic Condition Data Warehouse Home Page. Retrieved from https://www.ccwdata.org/web/guest/home

Chronic Condition Data Warehouse. (2018b). Condition Categories. Retrieved from https://www.ccwdata.org/web/guest/condition-categories

Cooper C, Tandy AR, Balamurali TB, & Livingston G (2010). A systematic review and meta-analysis of ethnic differences in use of dementia treatment, care, and research. Am J Geriatr Psychiatry, 18(3), 193–203. doi:10.1097/JGP.0b013e3181bf9caf [PubMed: 20224516]

Fisher TJ, Schwartz AC, Greenspan HN, & Heinrich TW (2016). Dementia: A complex disease with multiple etiologies and multiple treatments. Int J Psychiatry Med, 51(2), 171–181. doi:10.1177/0091217416636579 [PubMed: 26941207]

Ford E, Carroll JA, Smith HE, Scott D, & Cassell JA (2016). Extracting information from the text of electronic medical records to improve case detection: a systematic review. Journal of the American Medical Informatics Association : JAMIA, 23, 1007–1015. doi:10.1093/jamia/ocv180 [PubMed: 26911811]

Gitlin LN, Arthur P, Piersol C, Hessels V, Wu SS, Dai Y, & Mann WC (2018). Targeting Behavioral Symptoms and Functional Decline in Dementia: A Randomized Clinical Trial. J Am Geriatr Soc, 66(2), 339–345. doi:10.1111/jgs.15194 [PubMed: 29192967]

Gitlin LN, Kales HC, Marx K, Stanislawski B, & Lyketsos C (2017). A randomized trial of a web-based platform to help families manage dementia-related behavioral symptoms: The WeCareAdvisor. Contemp Clin Trials, 62, 27–36. doi:10.1016/j.cct.2017.08.001 [PubMed: 28800895]

Hane CA, Nori VS, Crown WH, Sanghavi DM, & Bleicher P (2020). Predicting Onset of Dementia Using Clinical Notes and Machine Learning: Case-Control Study. JMIR Med Inform, 8(6), e17819. doi:10.2196/17819 [PubMed: 32490841]

Hassanpour S, & Langlotz CP (2016). Information extraction from multi-institutional radiology reports. Artificial Intelligence in Medicine, 66, 29–39. doi:10.1016/j.artmed.2015.09.007 [PubMed: 26481140]

Hinton L, Franz C, & Friend J (2004). Pathways to dementia diagnosis: evidence for cross-ethnic differences. Alzheimer Dis Assoc Disord, 18(3), 134–144. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15494619 [PubMed: 15494619]

Kinatukara S, Rosati RJ, & Huang L (2005). Assessment of OASIS reliability and validity using several methodological approaches. Home Health Care Serv Q, 24(3), 23–38. doi:10.1300/J027v24n03_02 [PubMed: 16203688]

Lin PJ, Daly AT, Olchanski N, Cohen JT, Neumann PJ, Faul JD, … Freund KM (2021). Dementia Diagnosis Disparities by Race and Ethnicity. Med Care, 59(8), 679–686. doi:10.1097/MLR.0000000000001577 [PubMed: 34091580]

Lin PJ, Fillit HM, Cohen JT, & Neumann PJ (2013). Potentially avoidable hospitalizations among Medicare beneficiaries with Alzheimer's disease and related disorders. Alzheimers Dement, 9(1), 30–38. doi:10.1016/j.jalz.2012.11.002 [PubMed: 23305822]

Maclagan LC, Abdalla M, Harris DA, Stukel TA, Chen B, Candido E, … Bronskill SE (2023). Can Patients with Dementia Be Identified in Primary Care Electronic Medical Records Using Natural Language Processing? J Healthc Inform Res, 7(1), 42–58. doi:10.1007/s41666-023-00125-6 [PubMed: 36910911]

Matthews KA, Xu W, Gaglioti AH, Holt JB, Croft JB, Mack D, & McGuire LC (2019). Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged >/=65 years. Alzheimers Dement, 15(1), 17–24. doi:10.1016/j.jalz.2018.06.3063 [PubMed: 30243772]

Meystre SM, Savova GK, Kipper-Schuler KC, & Hurdle JF (2008). Extracting information from textual documents in the electronic health record: a review of recent research. Yearbook of medical informatics, 128–144. [PubMed: 18660887]

Noori A, Magdamo C, Liu X, Tyagi T, Li Z, Kondepudi A, … Das S (2022). Development and Evaluation of a Natural Language Processing Annotation Tool to Facilitate Phenotyping of Cognitive Status in Electronic Health Records: Diagnostic Study. J Med Internet Res, 24(8), e40384. doi:10.2196/40384 [PubMed: 36040790]

Oh IY, Schindler SE, Ghoshal N, Lai AM, Payne PRO, & Gupta A (2023). Extraction of clinical phenotypes for Alzheimer's disease dementia from clinical notes using natural language processing. JAMIA Open, 6(1), ooad014. doi:10.1093/jamiaopen/ooad014 [PubMed: 36844369]

Parker M, Barlow S, Hoe J, & Aitken L (2020). Persistent barriers and facilitators to seeking help for a dementia diagnosis: a systematic review of 30 years of the perspectives of carers and people with dementia. Int Psychogeriatr, 1–24. doi:10.1017/S1041610219002229

Pradeep KS, Sunder G, Bendale P, Mantri M, & Dande A (2013). Unified Medical Language System. In Pradeep KS, Sunder G, Bendale P, Mantri M, & Dande A (Eds.), Electronic Health Record: Standards, Coding Systems, Frameworks, and Infrastructures (pp. 145–152): Wiley-IEEE Press.

Romagnoli KM, Handler SM, & Hochheiser H (2013). Home care: more than just a visiting nurse. BMJ Qual Saf, 22(12), 972–974. doi:10.1136/bmjqs-2013-002339

Ryvicker M, Barron Y, Shah S, Moore SM, Noble JM, Bowles KH, & Merrill J (2022). Clinical and Demographic Profiles of Home Care Patients With Alzheimer's Disease and Related Dementias: Implications for Information Transfer Across Care Settings. J Appl Gerontol, 41(2), 534–544. doi:10.1177/0733464821999225 [PubMed: 33749369]

Samus QM, Johnston D, Black BS, Hess E, Lyman C, Vavilikolanu A, … Lyketsos CG (2014). A multidimensional home-based care coordination intervention for elders with memory disorders: the maximizing independence at home (MIND) pilot randomized trial. Am J Geriatr Psychiatry, 22(4), 398–414. doi:10.1016/j.jagp.2013.12.175 [PubMed: 24502822]

Sloane PD, Schifeling CH, Beeber AS, Ward KT, Reed D, Gwyther LP, … Zimmerman S (2017). New or Worsening Symptoms and Signs in Community-Dwelling Persons with Dementia: Incidence and Relation to Use of Acute Medical Services. J Am Geriatr Soc, 65(4), 808–814. doi:10.1111/jgs.14672 [PubMed: 28152160]

Speechly CM, Bridges-Webb C, & Passmore E (2008). The pathway to dementia diagnosis. Med J Aust, 189(9), 487–489. doi:10.5694/j.1326-5377.2008.tb02140.x [PubMed: 18976188]

Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, … Phelps CH (2011). Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement, 7(3), 280–292. doi:10.1016/j.jalz.2011.03.003 [PubMed: 21514248]

Stansfeld J, Crellin N, Orrell M, Wenborn J, Charlesworth G, & Vernooij-Dassen M (2018). Factors related to sense of competence in family caregivers of people living with dementia in the community: a narrative synthesis. Int Psychogeriatr, 1–15. doi:10.1017/S1041610218001394

Stearns MQ, Price C, Spackman KA, & Wang AY (2001). SNOMED clinical terms: overview of the development process and project status. Proc AMIA Symp, 662–666. Retrieved from https://pubmed.ncbi.nlm.nih.gov/11825268 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2243297/ [PubMed: 11825268]

The R Foundation. (2023). The R Project for Statistical Computing. Retrieved from https://www.r-project.org/

Topaz M, Adams V, Wilson P, Woo K, & Ryvicker M (2020). Free-Text Documentation of Dementia Symptoms in Home Healthcare: A Natural Language Processing Study. Gerontol Geriatr Med, 6, 2333721420959861. doi:10.1177/2333721420959861

Topaz M, Murga L, Bar-Bachar O, McDonald M, & Bowles K (2019). NimbleMiner: An Open-Source Nursing-Sensitive Natural Language Processing System Based on Word Embedding. Comput Inform Nurs, 37(11), 583–590. doi:10.1097/CIN.0000000000000557 [PubMed: 31478922]

Tullai-McGuinness S, Madigan EA, & Fortinsky RH (2009). Validity testing the Outcomes and Assessment Information Set (OASIS). Home Health Care Serv Q, 28(1), 45–57. doi:10.1080/01621420802716206 [PubMed: 19266370]

Werner P, Friedland RP, & Inzelberg R (2015). Alzheimer's disease and the elderly in Israel: are we paying enough attention to the topic in the Arab population? Am J Alzheimers Dis Other Demen, 30(5), 448–453. doi:10.1177/1533317515577130 [PubMed: 25794510]

References Cited in Appendix

Chae S, Song J, Ojo M, & Topaz M (2021). Identifying Heart Failure Symptoms and Poor Self-Management in Home Healthcare: A Natural Language Processing Study. Stud Health Technol Inform, 284, 15–19. doi:10.3233/SHTI210653 [PubMed: 34920459]

Koleck TA, Tatonetti NP, Bakken S, Mitha S, Henderson MM, George M, … Topaz M (2021). Identifying Symptom Information in Clinical Notes Using Natural Language Processing. Nurs Res, 70(3), 173–183. doi:10.1097/NNR.0000000000000488 [PubMed: 33196504]

Mikolov T, Sutskever I, Chen K, Corrado G, & Dean J (2013). Distributed representations of words and phrases and their compositionality. Advances in neural information processing systems. . Retrieved from https://papers.nips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf

Pradeep KS, Gaur S, Prashant B, Manisha M, & Atreya D (2013). Unified Medical Language System. In Electronic Health Record: Standards, Coding Systems, Frameworks, and Infrastructures (pp. 145–152): IEEE.

Quan H, Li B, Saunders LD, Parsons GA, Nilsson CI, Alibhai A, … Investigators, I. (2008). Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. Health services research, 43(4), 1424–1441. doi:10.1111/j.1475-6773.2007.00822.x [PubMed: 18756617]

Song J, Ojo M, Bowles KH, McDonald MV, Cato K, Rossetti SC, … Topaz M (2022). Detecting Language Associated With Home Healthcare Patient's Risk for Hospitalization and Emergency Department Visit. Nurs Res, 71(4), 285–294. doi:10.1097/NNR.0000000000000586 [PubMed: 35171126]

Stearns MQ, Price C, Spackman KA, & Wang AY (2001). SNOMED clinical terms: overview of the development process and project status. Proc AMIA Symp, 662–666. Retrieved from https://pubmed.ncbi.nlm.nih.gov/11825268 [PubMed: 11825268]

Topaz M, Lai K, Dowding D, Lei VJ, Zisberg A, Bowles KH, & Zhou L (2016). Automated identification of wound information in clinical notes of patients with heart diseases: Developing and validating a natural language processing application. Int J Nurs Stud, 64, 25–31. doi:10.1016/j.ijnurstu.2016.09.013 [PubMed: 27668855]

Topaz M, Murga L, Bar-Bachar O, McDonald M, & Bowles K (2019). NimbleMiner: An Open-Source Nursing-Sensitive Natural Language Processing System Based on Word Embedding. Comput Inform Nurs, 37(11), 583–590. doi:10.1097/CIN.0000000000000557 [PubMed: 31478922]

Topaz M, Murga L, Gaddis KM, McDonald MV, Bar-Bachar O, Goldberg Y, & Bowles KH (2019). Mining fall-related information in clinical notes: Comparison of rule-based and novel word embedding-based machine learning approaches. J Biomed Inform, 90, 103103. doi:10.1016/j.jbi.2019.103103 [PubMed: 30639392]

Topaz M, Woo K, Ryvicker M, Zolnoori M, & Cato K (2020). Home Healthcare Clinical Notes Predict Patient Hospitalization and Emergency Department Visits. Nurs Res, 69(6), 448–454. doi:10.1097/NNR.0000000000000470 [PubMed: 32852359]

Wake M, & Coenen A (1998). Nursing Diagnosis in the International Classification for Nursing Practice (ICNP). International Journal of Nursing Terminologies and Classifications, 9(4), 111–118. doi:10.1111/j.1744-618X.1998.tb00175.x

What this paper adds:

- This study validated a natural language processing (NLP) system specially developed for the home health care setting to identify signs and symptoms of Alzheimer's Disease and related dementias (ADRD) among home care patients.

- The paper demonstrates how, among patients admitted to home care without an established ADRD diagnosis, the ADRD signs and symptoms detected via NLP can help to identify patients in need of evaluation for ADRD.

Application of study findings:

- Home care clinicians are uniquely positioned to observe ADRD-related signs and symptoms which are evident in the home but less easily observed in other healthcare settings, and potentially to identify needs for specialist referrals (e.g. neurology, psychiatry, geriatrics), home safety interventions, and caregiver support services early in the ADRD disease trajectory.

- Applying NLP to clinicians' free-text clinical notes can help to identify patients at heightened risk for an eventual ADRD diagnosis and address unmet needs for these patients and their caregivers.

NLP Algorithm Creation

Step 1. Identifying concepts related to Alzheimer's disease and related dementias (ADRD)

Step 2. Creating a preliminary list of terms and expressions for each subcategory

Step 3. Creating language model: word embedding model (Word2Vec)

Step 4. Implementing an interactive vocabulary explorer to identify synonymous expressions

Evaluation of NLP algorithm

1 TRUE
2 FALSE
3 TRUE

Step 5. Evaluation of NLP algorithm performance: Label assignment on entire notes and review selected notes, then calculate precision
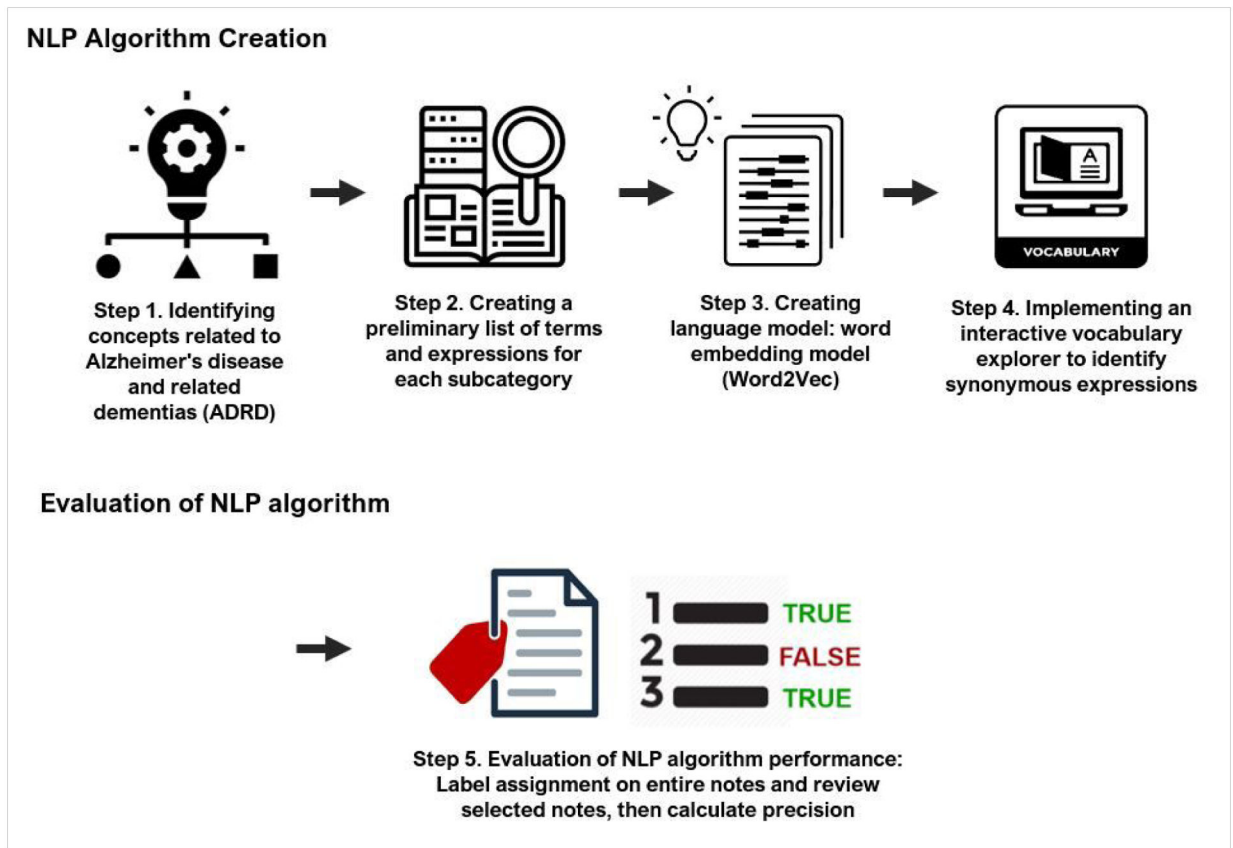
Figure 1.
Five step workflow for developing and evaluating the NLP algorithms
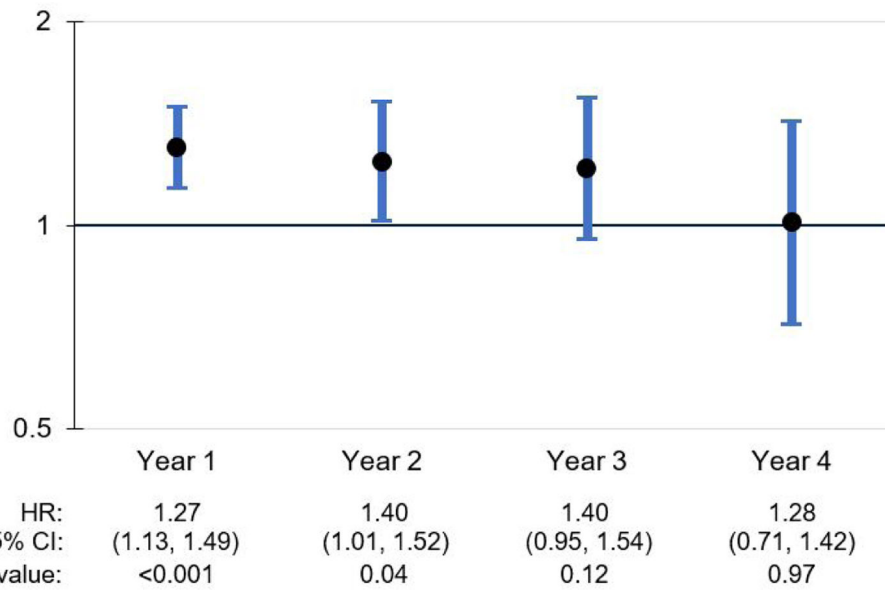
Figure 2.
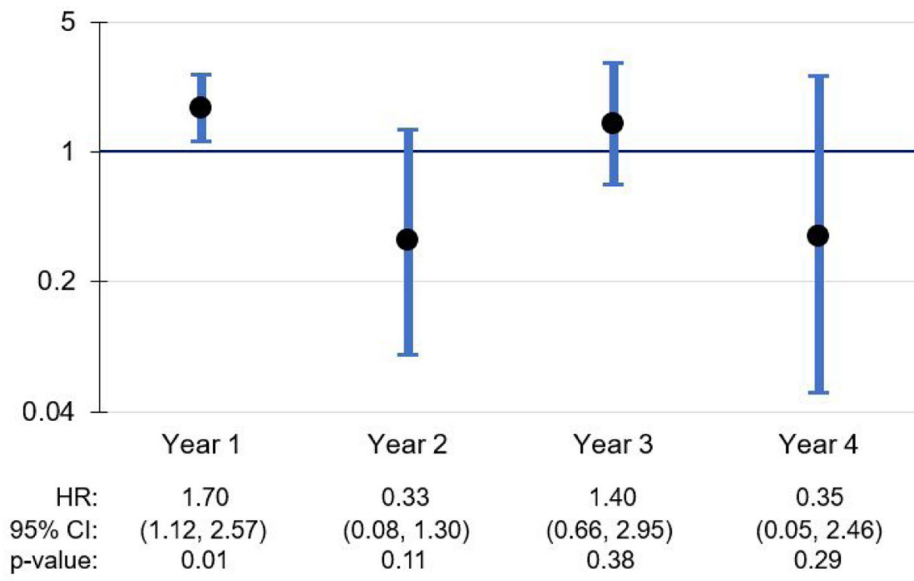ADRD mentions: Time-variant effects on ADRD diagnosis during follow-up period

Figure 3.
Caregiver stress: Time-variant effects on ADRD diagnosis during follow-up period
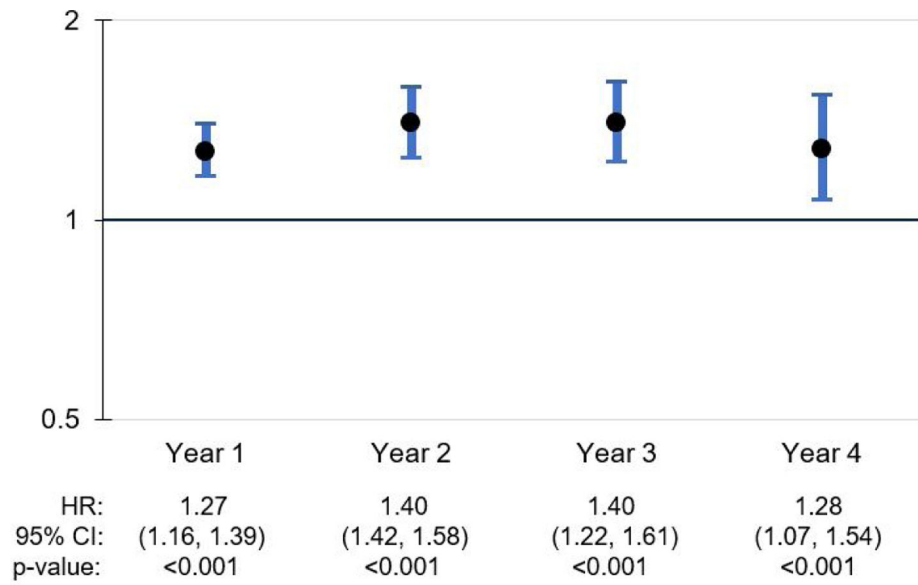
Figure 4.
Memory loss: Time-variant effects on ADRD diagnosis during follow-up period

Table 1.

Characteristics of overall patient sample and by ADRD diagnosis during 4-year follow-up period (N = 24,874)

	Overall Sample	ADRD Diagnosis During Follow-up	
		No	Yes
Characteristics	**N = 24,874**	**N = 15,973**	**N = 8,901**
Age at start of care: mean (SD)	76.6 (11.7)	73.5 (11.8)	82.2 (9.3)
Sex (female vs. male)	65.2%	64.7%	66.0%
Race			
Non-Hispanic White	66.4%	65.3%	68.4%
Black or African American	15.3%	15.6%	14.8%
Other	2.0%	2.1%	1.7%
Asian/Pacific Islander	5.0%	5.1%	4.7%
Hispanic	11.4%	11.9%	10.5%
Lives alone	46.6%	45.5%	48.5%
Type of impatient facility from which patient was discharged in 14 days prior to home care admission			
Short-stay acute hospital	57.3%	62.2%	48.4%
Inpatient rehabilitation facility	8.7%	9.1%	7.9%
Skilled nursing facility	8.4%	8.3%	8.4%
Long-term nursing facility, long-term care hospital, psychiatric hospital/unit, or other inpatient facility	2.9%	2.7%	3.2%
Not discharged from an inpatient facility	22.8%	17.6%	32.1%
Prior condition: impaired decision-making	3.6%	2.6%	5.4%
Prior condition: intractable pain	9.9%	11.5%	7.0%
Prior condition: memory loss	1.7%	0.8%	3.5%
Prior condition: urinary incontinence	12.9%	9.7%	18.6%
Hospitalization risk factors			
Recent mental decline	9.4%	7.5%	12.9%
Frailty	25.5%	22.8%	30.2%
History of falls	14.7%	11.3%	20.8%
Urinary incontinence or urinary catheter	30.8%	24.5%	42.1%
Bowel incontinence frequency			
Never or very rarely	93.0%	95.4%	88.6%
Less than once a week or more often	7.0%	4.6%	11.4%
Cognitive function			
Alert/oriented	79.0%	85.4%	67.5%
Requires prompting	18.2%	13.1%	27.3%
Requires assistance, considerable assistance or totally dependent	2.8%	1.5%	5.3%
When is the patient confused?			
Never, in new situations, on awakening, or at night only	98.2%	99.3%	96.0%
During the day and evening, or constantly	1.8%	0.7%	4.0%

	Overall Sample	ADRD Diagnosis During Follow-up	
		No	Yes
Characteristics	**N = 24,874**	**N = 15,973**	**N = 8,901**
Cognitive/behavioral/psychiatric symptoms:			
Impaired decision-making	5.5%	4.5%	7.3%
Memory deficit	2.5%	1.3%	4.8%
Overall prognosis: fragile health or serious progressive conditions that could lead to death with one year	14.8%	12.5%	19.0%
Hearing impairment	19.6%	13.8%	30.0%
Understanding of verbal content			
Clear comprehension without cues or repetitions	82.3%	87.6%	72.7%
Understands most conversations, but misses some parts/intent of message	16.4%	11.6%	24.9%
Only basic, rarely, or never	1.3%	0.8%	2.4%
ADL dependencies			
Grooming	27.7%	24.0%	34.4%
Dress upper body	41.4%	37.4%	48.5%
Dress lower body	67.3%	66.1%	69.5%
Bathing	87.4%	86.1%	89.7%
Toilet transferring	47.8%	43.9%	54.8%
Toilet hygiene	22.4%	18.8%	28.9%
Transferring	10.4%	8.3%	14.1%
Ambulation/Locomotion	29.2%	31.6%	24.7%

Table 2.

Timing of ADRD diagnosis during 4-year follow-up period (N = 24,874)

Timing of Diagnosis (Dx)	No Dx	Year 1	Year 2	Year 3	Year 4
N	15,973	3,555	2,254	1,834	1,258
% of overall sample	64.2%	14.3%	9.1%	7.4%	5.1%
% of patients with Dx (N=8,901)		40.0%	25.3%	20.6%	14.1%

Table 3.

Prevalence and examples ADRD constructs in free-text clinical notes (N = 24,874)

Domains & Constructs	Overall Sample	ADRD Dx During Follow-Up		Example of Construct in Free-Text Notes
		No	Yes	
	N = 24,874	N = 15,973	N = 8,901	
ADRD disease names				
Mentions of ADRD	3.1%	1.9%	5.1%	Pt (patient) on chair hx (history) of last stage *dementia* [ADRD disease names] unable to answer questions
Symptoms				
Memory loss	10.6%	6.7%	17.6%	Pt was at times tearful today is *extremely forgetful* [Memory loss] will ask the same question many times
Unspecified cognitive deficits	2.2%	1.4%	3.8%	Pt is *cognitively impaired* [Unspecified cognitive deficits] and does not initiate any conversation
Communication problems	3.9%	2.9%	5.5%	patient with significant *word-finding difficulties, phonemic paraphasias, and verbal apraxia* [Communication problems] is on speech tx (therapy)
Motor issues	11.0%	9.0%	14.7%	pt is homebound due to *unsteady gait* [Motor issue] and steep stairs
Physically/verbally aggressive	6.5%	5.6%	8.1%	pt combative refused vital signs *yelled at vn* (visiting nurse) [Physically aggressive] don't touch me *attempted to hit vn* [Verbally aggressive]
Poor sleep	2.7%	2.3%	3.3%	pt reports having *trouble sleeping* [Poor sleep] pt states he spends his nights thinking about many things
Wandering	0.2%	0.1%	0.3%	pt continues to *wander* [Wandering] day and night trying to get out of apt
Disorientation	0.5%	0.2%	0.9%	the client *became disoriented* [Disorientation] and did not recognize family nor her own location
Hoarding and clutter	4.1%	3.5%	5.2%	pt's surroundings are very crowded and pt appears to *collect and hoard* [Hoarding and clutter] with stacks of papers and magazines everywhere
Uncooperative	7.2%	5.6%	10.1%	pt *refused to* [Uncooperative/ refusal] learn how to self-inject insulin
Depression or apathy	11.4%	10.0%	14.0%	patient continues to be fatigued reports that hair is falling down and *feels depress* [Depression/ apathy] from fighting cancer
Anxiety	19.6%	18.1%	22.4%	Pt *very anxious* [Anxiety] needs ongoing instruction on disease management
Delusions or hallucinations	0.7%	0.4%	1.2%	pt continues to respond verbally to *auditory hallucinations* [Delusions/hallucinations]
Agitation	2.8%	2.0%	4.3%	pt *very agitated* [Agitation] during vn visit
Confusion	4.4%	2.6%	7.5%	pt continues to cry every afternoon *more confused* [Confusion] in evening
Bladder or bowel incontinence	3.2%	2.3%	4.7%	pt is alert and orientedx3 *incontinent in bowel and bladder* [Bladder/bowel incontinence] skin is intact
Weight loss	2.7%	2.2%	3.5%	pt reports 15lbs *wt (weight) loss* [Weight loss] x 6wks after her husband passed away
Poor/increased appetite	2.7%	2.1%	3.6%	pt still co of abdominal bloating has *no appetite* [Poor/increased appetite] went to gastroenterologist last week

Domains & Constructs	Overall Sample	ADRD Dx During Follow-Up		Example of Construct in Free-Text Notes
		No	Yes	
	N = 24,874	N = 15,973	N = 8,901	
Trouble eating or swallowing	1.8%	1.3%	2.7%	pt *unable to swallow* [Trouble eating or swallowing] Ceftin tabs due to large size pharmacy contacted and Ceftin confirmed to be ok to crush and mixed with pudding
Executive function problems	1.5%	1.1%	2.3%	Patient *needs constant verbal cues* [Executive function problems] for safety and proper techniques
Functional issues (ADL/IADLs)	0.4%	0.4%	0.5%	pt having increased *difficulty feeding self* [Functional issues] suggested weighted utensils
Hearing loss	1.3%	0.7%	2.3%	pt lacks spontaneity and alertness due to *hearing deficits* [Hearing loss]
Caregiver stress	0.2%	0.2%	0.4%	pts *wife overwhelmed* [Caregiver stress] and comfort provided
Interventions				
Addressing unmet needs	0.7%	0.7%	0.7%	educate pts *son on how to use* [Indication of unmet need] a glucometer
Referrals for end-of-life care	3.1%	2.3%	4.4%	pts md will discuss with pt possible *transfer to hospice* [Referrals for end-of-life care]
ADRD medications	2.5%	1.9%	3.6%	Daughter to pick up *aricept* [Specific medications] from pharmacy
Home safety interventions	57.7%	54.6%	63.3%	*grab bar for toilet* [Home safety interventions] ordered

Table 4.

Multivariate Cox proportional hazards model with time-varying coefficients: Associations of ADRD-related constructs with ADRD diagnosis during follow-up (N = 24,874)

ADRD Diagnosis	Year 1	Year 2	Year 3	Year 4
	HR [95% CI]	HR [95% CI]	HR [95% CI]	HR [95% CI]
ADRD mentions	1.297** [1.127–1.493]	1.24* [1.011–1.52]	1.211 [0.949–1.544]	1.006 [0.714–1.416]
Memory loss	1.273** [1.164–1.394]	1.398** [1.24–1.578]	1.401** [1.218–1.611]	1.278* [1.065–1.535]
Motor issues¥	1.102* [1.036–1.171]			
Hoarding and clutter¥	1.139* [1.033–1.255]			
Uncooperative¥	1.235** [1.148–1.33]			
Delusions or hallucinations¥	1.252* [1.022–1.535]			
Agitation	1.337** [1.145–1.56]	1.18 [0.938–1.484]	1.004 [0.754–1.337]	1.8** [1.334–2.428]
Confusion	1.397** [1.24–1.575]	1.022 [0.845–1.235]	1.048 [0.84–1.307]	1.115 [0.846–1.47]
Poor/increased appetite	1.286* [1.091–1.517]	1.305* [1.038–1.642]	0.925 [0.678–1.262]	0.649 [0.415–1.013]
Executive function problems	1.173 [0.961–1.433]	1.212 [0.907–1.62]	1.015 [0.698–1.475]	0.523* [0.286–0.954]
Caregiver stress	1.695* [1.118–2.57]	0.325 [0.081–1.304]	1.399 [0.663–2.952]	0.345 [0.048–2.459]
Referrals for end-of-life care	1.408** [1.213–1.634]	1.193 [0.96–1.481]	0.909 [0.685–1.207]	1.191 [0.873–1.626]
ADRD-specific medications	1.258* [1.065–1.487]	1.206 [0.949–1.532]	1.129 [0.842–1.513]	1.009 [0.683–1.491]

Note: HR=hazard ratio. CI=confidence interval.

*
p<0.05.

**
p<0.001.

¥Variable met the proportionality assumption and therefore did not have time-varying coefficients.