



Published in final edited form as:

J Biomed Inform. 2021 October ; 122: 103903. doi:10.1016/j.jbi.2021.103903.

ReHouSED: A Novel Measurement of Veteran Housing Stability Using Natural Language Processing

Alec B Chapman^{1,2}, Audrey Jones^{1,2}, A. Taylor Kelley^{1,3}, Barbara Jones^{1,4}, Lori Gawron^{1,5}, Ann Elizabeth Montgomery^{6,7,8}, Thomas Byrne^{8,9,10}, Ying Suo^{1,2}, James Cook^{1,2}, Warren Pettey^{1,2}, Kelly Peterson^{1,2,11}, Makoto Jones^{1,2}, Richard Nelson^{1,2,8}

¹Informatics, Decision-Enhancement and Analytic Sciences (IDEAS) Center, Veterans Affairs (VA) Salt Lake City Health Care System, Salt Lake City, UT

²Division of Epidemiology, University of Utah School of Medicine, Salt Lake City, UT

³Division of General Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT

⁴Division of Pulmonary and Critical Care Medicine, University of Utah and VA Healthcare System, Salt Lake City, UT

⁵Department of Obstetrics and Gynecology, University of Utah School of Medicine, Salt Lake City, UT

⁶Birmingham Veterans Administration Medical Center, Birmingham, AL

⁷School of Public Health, University of Alabama at Birmingham, Birmingham, AL

⁸U.S. Department of Veterans Affairs, National Center on Homelessness among Veterans, Tampa, FL

⁹U.S. Department of Veterans Affairs, Center for Healthcare Outcomes and Implementation Research, Edith Nourse Rodgers VA Medical Center, Bedford, MA

¹⁰Boston University School of Social Work, Boston, MA

¹¹Veterans Health Administration Office of Analytics and Performance Integration

Abstract

Housing stability is an important determinant of health. The US Department of Veterans Affairs (VA) administers several programs to assist Veterans experiencing unstable housing. Measuring long-term housing stability of Veterans who receive assistance from VA is difficult due to a lack of standardized structured documentation in the Electronic Health Record (EHR). However, the text of clinical notes often contains detailed information about Veterans' housing situations that may be extracted using natural language processing (NLP). We present a novel NLP-based measurement of Veteran housing stability: Relative Housing Stability in Electronic Documentation (ReHouSED). We first develop and evaluate a system for classifying documents containing information about Veterans' housing situations. Next, we aggregate information from multiple documents to derive a patient-level measurement of housing stability. Finally, we demonstrate

this method's ability to differentiate between Veterans who are stably and unstably housed. Thus, ReHouSED provides an important methodological framework for the study of long-term housing stability among Veterans receiving housing assistance.

Keywords

Natural Language Processing; Homelessness; Veterans Affairs; Information Extraction; Social Determinants of Health

1. Introduction

Homelessness and lack of access to stable housing represent a major public health crisis, resulting in poor health outcomes and high treatment costs [1–3]. The US Department of Veterans Affairs (VA) administers several nationwide programs for addressing homelessness among Veterans. One such program is the Supportive Services for Veteran Families (SSVF) program, which partners with non-profit organizations throughout the country to provide services to help Veterans at risk of homelessness to maintain their housing and to assist those who are currently experiencing homelessness to gain stable housing as quickly as possible. SSVF provides an array of flexible services to help Veterans maintain or obtain stable housing, including outreach, case management, and temporary financial assistance.

A Veteran's housing situation is not consistently recorded in structured Electronic Health Record (EHR) data maintained by VA. Much of the information relevant to housing is instead documented in unstructured free-text notes. Previous studies have shown that evidence of housing instability can be extracted from clinical text using natural language processing (NLP)[4–7]. However, determining a Veteran's precise housing circumstance at any single point in time remains challenging.

Housing is a complex variable which changes as Veterans move in and out of different housing circumstances (e.g., staying with a friend, sleeping in a car, obtaining an apartment). Existing NLP systems were not designed to capture changes in Veteran housing status, such as determining when housing is obtained. Additionally, these systems do not attempt to infer a patient-level housing situation and are limited to identifying homelessness within individual clinical texts which may contain outdated or inconsistent information. These limitations restrict the longitudinal study of a Veteran's housing stability, which is necessary for assessing long-term outcomes and evaluating the effectiveness of programs like SSVF.

In this work, we present an NLP system for identifying the current housing status of Veterans who were previously homeless. First, we annotated a corpus of clinical notes belonging to Veterans who had received assistance from SSVF. We then developed an NLP system for classifying whether a Veteran is stably housed based on evidence in an individual note. Following validation of this system, we processed notes for a cohort of Veterans over a period of several months. We then use the NLP-derived document classifications to calculate a novel measure of patient-level housing stability which we call Relative Housing Stability in Electronic Documentation (*ReHouSED*). We demonstrate the capability of this method to

distinguish between varying levels of housing instability at different points in time and show its accuracy compared to standard diagnosis coding used for medical billing purposes.

2. Background

Previous studies have found that individuals experiencing homelessness are at increased risk for adverse physical and mental health outcomes, as well as mortality [8,9]. Veterans have historically been over-represented in the homeless population [10]. In 2009, VA made preventing and ending homelessness among Veterans an explicit policy goal, cutting the number of homeless Veterans by half over the next decade. During this time, VA implemented and expanded several programs which aim to reduce Veteran homelessness; the two largest of these are the U.S. Department of Housing and Urban Development-VA Supportive Housing (HUD-VASH) program and the SSVF program. HUD-VASH provides Veterans who have experienced homelessness with ongoing housing subsidies matched with supportive services intended to help them maintain stable housing. SSVF awards grants to community-based agencies who in turn provide housing support for Veterans who are either currently homeless (“rapid rehousing”) or at risk of becoming homeless (“homelessness prevention”). One component of SSVF is temporary financial assistance (TFA), which provides direct financial support for Veterans to pay housing-related costs.

Previous studies have examined the impact of these programs on improving housing gains and health outcomes for homeless Veterans. One study provided evidence that the expansion of HUD-VASH was a key driver of declines in Veteran homelessness nationwide between 2007 and 2017 [11]. Nelson et al. found that VA’s HUD-VASH program was associated with decreased use of Medicare services [12]. Another study showed that 81.4% of Veterans enrolled in SSVF obtained housing by program discharge, as well as an association between receiving TFA and being housed upon program discharge [13]. While there is evidence of links between VA housing programs and Veterans’ entry into housing, the impact of these programs on long-term housing stability is relatively unknown due to challenges of documentation in EHRs.

There are limitations to established methods for determining housing instability in EHRs. For instance, International Statistical Classification of Diseases and Related Health Problems (ICD) diagnostic codes have been shown to be problematic at accurately identifying patients with housing problems due to a lack of specificity and rapid shifts between different housing situations [14]. Housing stability may change rapidly and frequently, creating problems for EHR data fields that are copied from one note to another and infrequently updated. Furthermore, there is currently no ICD code indicating that an individual has exited homelessness and is currently stably housed. Another method for determining homelessness is the use of NLP on free-text narratives, where housing status and other social risk factors are often recorded [4–7]. Prior NLP studies focused primarily on extracting evidence of homelessness at the note level but not evidence of an individual’s transition to stable housing. Because a Veteran’s medical record may contain conflicting information about their housing situation, determining housing status requires reconciling the information present in multiple notes at different points of time. To our knowledge, no

previous study has attempted to extract patient-level housing status from multiple clinical documents.

Our research aims to address this need by developing an NLP system to determine patient-level housing status and detect changes in housing stability over time. The system we describe could have many practical applications for future research and operations including evaluating the long-term housing outcomes associated with housing assistance programs, both within and outside VA. Our method of aggregating information extracted from multiple clinical documents with NLP could also be applied to other clinical domains to construct patient-level classifications of time-varying constructs.

3. Methods

At a high level, our approach to measuring patient-level housing stability consisted of three steps. First, we retrieved all free-text EHR documents for a set of Veterans containing relevant keyword hits from VA Corporate Data Warehouse (CDW) and organized them in batches of 30-day intervals. Second, each document was processed by a rule-based NLP system to obtain a classification of the Veteran's housing status based on that individual document. Third, the predicted document classifications in each time window were aggregated and used to calculate a patient-level measurement of housing stability (*ReHouSED*). This process is shown in Figure 1.

We validated our system at two different levels. First, we measured our system's performance at classifying individual documents using a reference standard set of annotated notes from SSVF participants. Second, we measured our system's ability to generate patient-level housing status by comparing it against structured data recorded as part of the SSVF program.

3.1 Veteran Population and Document Selection

Veterans enrolled in the SSVF program were identified using SSVF administrative data. These data are collected by SSVF grantees via Homeless Management Information Systems (HMIS), which are used to record standardized client-level information on characteristics of individuals experiencing homelessness and accessing services through federally funded housing assistance programs¹. A Veteran's housing status immediately following discharge from the SSVF program is recorded as a structured data field in SSVF administrative data by the case manager assisting the Veteran. This manually assigned classification was used for document sampling and as a reference standard for assessing the accuracy of patient-level housing status.

We then linked Veterans found in HMIS to VA CDW to identify relevant clinical documents. Our Veteran population consisted of any Veteran who had enrolled in SSVF and had received assistance for rapid rehousing, implying homelessness at the time of entry to SSVF. Veterans may participate in one or more "episodes" of SSVF, defined as entry into the program and a subsequent exit. At the end of each episode, Veterans were recorded by SSVF

¹ <https://www.hudexchange.info/programs/hmis>

as exiting to a destination of “Stably Housed”, “Unstably Housed”, or “Not Housed”. For the purpose of our analysis, “Unstably Housed” and “Not Housed” episodes were grouped into a single category of “Unstably Housed”, which accounted for approximately 18% of the 95,022 SSVF episodes (Table 1).

We utilized data from VA CDW to retrieve free-text notes for annotation and classification. To identify relevant documents, we queried notes of SSVF participants containing keywords related to housing. The keywords used in our search are shown in Table A.1 in the Appendix. The majority of notes retrieved using this strategy primarily were either clinical documentation from medical visits or social work notes authored by VA case managers assigned to assist the Veteran with issues including housing.

3.2 Document Annotation

Whereas we were able to use the manually assigned housing status destination as the reference standard for Veteran patient-level housing status, there was no existing reference standard for document-level classifications. Therefore, we manually annotated a corpus of documents. To include documents with mentions of stable housing as well as documents with mentions of unstable housing, we block-sampled within strata defined by the Veteran’s destination (“Stably Housed” vs. “Unstably Housed”) and time period (up to 6 months before SSVF enrollment, between entry to SSVF and exit, and up to 6 months following exit from SSVF). A total of 621 notes were sampled using this sampling strategy, of which 257 were used for training, 155 for validation, and 209 were used as a blind testing set.

The final corpus of 621 notes was annotated by two domain expert annotators. For each document, the annotator was directed to assign one of three document labels: “**Stably Housed**”, “**Unstably Housed**”, and “**Unknown**”. Detailed definitions for each of the classes are shown in Table A.2 in the appendix.

An additional four batches of 90 notes each were initially sampled as practice batches to develop an annotation schema and to train annotators. Once an acceptable inter-annotator agreement (IAA) was achieved on practice batches (Cohen’s Kappa = 0.75) [15], the corpus of 621 notes was split into 10 batches. Each annotator independently annotated 5 batches. 70 documents were classified by both annotators and used to measure a final IAA.

3.3 Natural Language Processing System

We developed a rule-based system for extracting entities within a document and inferring a document-level classification. This NLP system was implemented using medspaCy [16], a toolkit for clinical NLP within the Python spaCy framework². medspaCy has previously been used in the VA for COVID-19 surveillance [17]. The code and knowledge base used by our system has been made publicly available³.

² <https://spacy.io/>

³ <https://github.com/abchapman93/ReHouSED>

As shown in Figure 1, the document-level NLP system includes four steps: 1) entity extraction; 2) attribute detection; 3) postprocessing; and 4) document classification. Each step is described below.

3.3.1 Entity Extraction—We defined an entity as a span of text within a document representing a housing-related concept. Building on previous work [5,6], we specified several entity classes related to homelessness. In addition, we added a class relating specifically to stable housing. Five entity classes were defined in total. Examples are shown in Table 2.

The vocabulary in our corpus was found to be highly variable and often ambiguous. As such, rules were designed to be generalizable while also attempting to resolve ambiguities. Rules consisted of both semantic patterns which matched exact phrases and syntactic patterns which matched varied spans of texts. A total of 381 rules were used to extract entities.

3.3.2 Attribute detection—Terms related to housing are often used in text without explicitly referring to the Veteran’s current housing status. For example, the provider or case manager who authored the note may be discussing the Veteran’s past housing issues, their goals regarding housing, or explicitly negating a housing-related concept. To differentiate these instances from entities describing a Veteran’s current housing status, our system assigned attributes for each entity based on whether the entity was negated, historical, or hypothetical. Additionally, ambiguous or irrelevant snippets of text from semi-structured templates, such as educational text or questionnaires, were explicitly set to be ignored. An entity was considered *asserted* if all four of these attributes were false. Table 3 shows examples of entities assigned each of these attributes.

Attributes for each entity were detected using several methods in the text processing pipeline. First, the ConText algorithm [18] was used to identify linguistic modifiers in the text surrounding a mention. Second, sections of the note were extracted using medspaCy’s section detection module. This allowed section-specific logic such as marking mentions of homelessness in “Past Medical History” as historical or mentions of housing in “Goals” as hypothetical.

3.3.3 Postprocessing—The final step of our NLP system was a postprocessing layer which combined the results of upstream components to resolve ambiguities and enforce additional constraints to prevent false positives. Postprocessing rules modified or removed entities based on properties such as exact text, attributes, and their surrounding context in the document. For example, we found that the terms “home” and “apartment” occur frequently without directly referring to a Veteran’s housing situation. A postprocessing rule was implemented to ignore these terms unless there was contextual information implying the entity was connected to the Veteran (e.g., “he lives in an apartment” or “she lost her home”). Another rule differentiated between instances of the term “rent” that refer either to evidence of stable housing (“he pays \$600/month for rent”) or housing risk (“she cannot pay her rent”).

A total of 34 postprocessing rules were used in our final system. These rules could sometimes cause false negatives. For example, a mention of “apartment” could result in a false negative if the note states elsewhere in the document (but not in the same sentence as the entity) that the Veteran lives in the apartment. Despite the risk of false negatives, during development it was found that postprocessing rules ultimately improved performance by reducing false positives.

3.3.4 Document Classification—The primary purpose of the NLP system is to assign a document-level classification of a Veteran’s housing status. Information about housing status in clinical texts is often vague or inconsistent. For example, a note may introduce a Veteran as a “homeless veteran” but later state that they have moved into stable housing. As such, making a document-level classification of a Veteran’s housing status requires extracting all mentions of housing-related concepts, detecting the linguistic attributes as described above, and then inferring the overall housing status based on all entities found in the text.

We developed a heuristic algorithm to assign a document-level classification to each note. The logic of this algorithm is shown in Figure 2. First, the document is parsed for specific note sections, such as “Housing Status”, or semi-structured questionnaires that contain clear, definitive documentation of the Veteran’s housing status. If there are no such sections, the document label is inferred based on entities found in other parts of the note following these steps:

- Predict **“Stably Housed”** if either:
 - There is at least one asserted mention of stable housing
 - There are no mentions of housing instability **and** there is an explicitly negated statement of homelessness (e.g., “The patient is not homeless”)
- Predict **“Unstably Housed”** if there is no evidence of stable housing **and** there is at least one:
 - Asserted mention of an “Unstable Housing” concept
 - Hypothetical mention of “Stable Housing” (e.g., “He is looking for stable housing”)
 - Otherwise, no relevant information has been identified and the predicted label is **“Unknown”**

3.3.5 Machine Learning Model Comparisons—We compared our rule-based NLP classifier with two machine learning classifiers. First, as a baseline we trained a bag-of-words (BOW) model using n-grams (n=1–3). Second, we developed a hybrid model (NLP+ML), which predicted a document classification based on features derived from the rule-based NLP model. Both machine learning models used the XGBoost classification algorithm [19]. We hypothesized that a machine learning model may learn a more sophisticated decision function than the heuristic NLP classifier, which was very strict and sensitive to entity-level errors. The feature set used with the hybrid model consisted

of n-grams as well as the counts of entity labels, attributes, section categories, and linked ConText modifiers.

3.4 Patient-Level Housing Status

3.4.1 Definition of Measurement of Patient-Level Housing Status—Our ultimate objective was to determine an overall patient-level classification of housing status. Following the classification of a set of documents for a Veteran, the next step was to aggregate the extracted information to a patient level. Determining a patient-level classification was complicated by the fact that Veterans could have several documents in a close window of time with conflicting information about their housing status. This variation may be due to inconsistent documentation, actual change in housing between visits, or NLP errors.

While there may be some variation between notes, we hypothesized that Veterans who were in stable housing during a period of time would have a higher proportion of notes documenting that they were stably housed. This statistic could then act as a measurement of overall housing stability and be used for patient- and population-level analysis.

We defined a patient-level measurement of housing stability as the proportion of documents classified by our NLP system as “Stably Housed” over a 30-day window (Eq. 1). We refer to this measurement as Relative Housing Stability in Electronic Documentation (*ReHouSED*). We limit the notes used in this measure to notes classified either as “Stably Housed” or “Unstably Housed”, excluding notes which were classified as “Unknown” or did not contain any relevant housing keywords.

$$\text{ReHouSED} = \frac{\text{Count (Stably Housed)}}{\text{Count (Stably Housed)} + \text{Count (Unstably Housed)}}$$

Eq 1. Formula for calculating a Veteran’s *ReHouSED* score over a 30-day time window.

3.4.2 Sample for Evaluating *ReHouSED* Measure—To evaluate this method, we compared *ReHouSED* scores for a sample of Veterans before and after treatment in SSVF. As a reference standard, we used the housing destination variable recorded in HMIS data upon a Veteran’s exit from the SSVF program. We assumed that a Veteran maintained the same housing for at least 30 days after discharge from SSVF.

To construct a sample of Veterans and documents for evaluating *ReHouSED*, we first identified a group of 10,328 Veterans from our testing set. Each Veteran had exactly 1 SSVF episode and had received assistance for rapid rehousing. We then retrieved notes for these Veterans from VA CDW using the keywords described previously. In order to compare text documents before and after SSVF treatment, we retrieved notes which were authored in one of two 30-day time intervals: 60–90 days before exiting SSVF (“pre-SSVF”) and 0–30 after (“post-SSVF”). Each of these notes were processed by the NLP system and assigned a document classification using the steps described above. Any document classified as “Unknown” was excluded from further analysis. Finally, we excluded any Veteran who

did not have at least one “Stably Housed” or “Unstably Housed” document in both the pre-SSVF and post-SSVF time intervals.

3.4.3 Analysis of *ReHouSED* Measure—*ReHouSED* scores calculated before and after SSVF were compared between the Veterans who exited the SSVF program to stable housing (N=3,279) and those who exited to unstable housing (N=415) to evaluate whether this metric could potentially be used to distinguish between these two housing situations. The median and interquartile range (IQR) of *ReHouSED* scores were calculated within each group. Mann-Whitney rank tests were performed in each cross-group comparison to assess statistical significance between stable and unstably housed groups both before and after SSVF.

3.4.4 Patient Housing Status Classification—To further validate the utility of the *ReHouSED* measure in characterizing housing stability, we compared it against ICD diagnostic codes at classifying episodes of unstable housing. Any post-SSVF episode with *ReHouSED* 0.4 was classified as “Stable”, while any episode below that threshold was classified as “Unstable”. This threshold was found empirically by calculating the true positive rate and false positive rate at various classification thresholds and choosing the value that maximized their geometric mean. For the ICD-10 code comparison, any Veteran who had at least one diagnosis code representing homelessness or housing instability during the 30-day episode window was considered “Unstably Housed”. Otherwise, they are considered “Stably Housed”. The list of ICD-9/10 codes are shown in Table A.3 in the appendix.

4. Results

We report results for the following three tasks: 1) annotation of a corpus of 621 sampled notes; 2) NLP document classification and entity extraction using the annotated corpus; and 3) analysis of patient-level *ReHouSED* scores of a large sample of Veterans.

4.1 Document Annotation

The final annotated corpus of 621 notes contained 139 (22%) documents classified as “Stably Housed”, 329 (53%) as “Unstably Housed”, and 153 (25%) as “Unknown”. We measured inter-annotator agreement on 70 documents across the training and test sets which were annotated by both annotators, achieving a Cohen’s Kappa of 0.7. All disagreements were adjudicated and assigned a final classification by consensus.

Most disagreements involved one annotator assigning a label of “Unknown” while the other assigned “Stably Housed” or “Unstably Housed”. These notes typically contained some mention of housing assistance or participation in a group session discussing housing but lacked a definitive statement of the Veteran’s current housing circumstance. Thus, the annotator had to judge whether to consider a phrase relevant evidence. Other disagreements involved uncommon phrasing (“domiciled with friend”), only contained information in half-completed template forms, or described a scenario which had not been encountered in the practice batches (e.g., impending release from prison).

4.2 Natural Language Processing System

4.2.1 Document Classification—We report the precision, recall, and F1 for document classification for the rule-based NLP classifier within each class, as well as macro-averaged scores. Results are shown in Table 4.

Overall, the NLP classifier achieved moderate performance in each class. F1 and precision were highest among “Unstably Housed” documents. The highest recall was achieved for “Unknown” documents, which exhibited low precision, showing that many of the “Stably Housed” and “Unstably Housed” documents were incorrectly classified as “Unknown”, which was an issue similarly encountered in annotation.

4.2.2 Machine Learning Comparisons—We then compared the rule-based NLP classifier to a BOW baseline and hybrid NLP + ML model. Macro-averaged scores are reported in Table 5. The NLP classifier significantly outperformed the BOW model, showing that the additional information extracted from NLP is more effective than just the raw text. The rule-based NLP classifier also performed slightly higher than the hybrid NLP + ML model in all metrics.

4.2.3 Error Analysis—We reviewed each of the 69 documents classified incorrectly by the rule-based classifier and identified the cause of the incorrect prediction. Each document was categorized into one of five error types. Specific examples and explanations are provided in Table A.4 in the appendix. The most frequent errors were caused by false negative mentions (n=23). Some of these missed concepts were absent from the lexicon, such as “deliver furniture to the Veteran” or “housing pursuit”. Other spans were correctly extracted from the text but incorrectly ignored due to postprocessing rules which were used to reduce false positives.

The second-most frequent cause of error was an incorrect entity attribute (n=16), such as incorrectly identifying an entity as hypothetical or failing to recognize that a concept is historical.

Due to the complexity of the task and the modest inter-annotator agreement, a number of documents could be reasonably assigned a different label than what was assigned by the annotator (n=16). Some instances were likely due to annotator error, but more frequently the information in a note was either unclear, contradictory, or implicitly stated, requiring some subjective judgment from the annotator.

The remaining errors fell into two categories. The first was false positives (n=10) caused by lexical ambiguity, incorrect parsing, or mentions which should have been ignored. The second was other issues such as poor text formatting or copy-and-pasted documentation which was out of date (n=4).

4.2.4 Entity-Level Analysis—The counts of the most frequent entity texts and labels in the entire annotated corpus are provided in the appendix in Table A.5. Terms often depended on the context around target terms to signify whether the text was informative regarding a Veteran’s current housing status, as well as whether the term referred to stable or unstable

housing. This ambiguity is demonstrated visually in Figure 3, which shows the relative frequency with which the most common terms occurred in each of the three document classes. Many terms occurred frequently in more than one document class and their presence in a document alone could not be used to classify a document. Some terms were found to be more strongly associated with a specific class, such as “shelter”, “home visit”, and “housing search”.

4.3 Patient-Level Housing Status

4.3.1 Sample for Evaluating *ReHouSED* Measure—Our final objective was to derive a patient-level measurement of Veteran housing stability using a measure called *ReHouSED*. The results of the sampling method to evaluate *ReHouSED* are summarized in Table 6. Of the 10,328 Veterans initially considered in our sample, 7,131 had a note in the relevant time periods, with a total of 62,028 documents to be processed by the NLP. After excluding “Unknown” documents and Veterans who did not have at least one document both pre- and post-SSVF, a total of 3,694 Veterans and 35,452 documents were included in our final sample.

4.3.2 Analysis of *ReHouSED* Measure—Comparing aggregate *ReHouSED* values by time period and housing destination showed a significant difference in housing stability which aligns with the destination data from the HMIS dataset. Detailed statistics are shown in Table 7. Figure 4 compares the distribution and median *ReHouSED* values by outcome over the two time intervals.

Veterans who were stably housed at the time of exit from SSVF tended to have a higher *ReHouSED* than those who were unstably housed (median=0.83 versus 0.11, $p<0.001$). Furthermore, it was found that the two groups showed significantly different pre-SSVF *ReHouSED* scores (median=0.25 versus 0.05, $p<0.001$). While the median of both stably and unstably housed Veterans increased after exiting SSVF, the increased *ReHouSED* score observed among stably housed Veterans was much higher than among unstably housed Veterans (increase=0.58 versus 0.11).

4.3.3 Patient Housing Status Classification—Our final evaluation was to use *ReHouSED* values as a binary classifier for detecting “Unstably Housed” episodes. Classification metrics for the *ReHouSED* classifier and ICD diagnosis codes are shown in Table 8. The *ReHouSED* episode classifier significantly outperformed the ICD-10 codes in all metrics, achieving an F1 of 44.6 and AUROC of 80.3. Both methods achieved much higher recall than precision. The largest difference between the two methods was in specificity, which corresponded to recall for detecting “Stably Housed” episodes.

To qualitatively understand the differences between these two methods, we sampled 25 Veterans from each destination where the NLP classification was consistent with the reference standard but differed from diagnosis codes. We reviewed notes and diagnosis codes assigned to each Veteran during the thirty-day period following exit from SSVF.

In all of the reviewed “Stably Housed” Veterans, one or more notes clearly documented that a Veteran had been housed but a diagnosis code indicating homelessness or housing

instability was assigned as a diagnosis for the visit. This was likely because the Veteran's progress on housing issues was discussed but there was no diagnosis code to indicate that the Veteran has exited homelessness. ICD codes were sometimes inserted into the text, (e.g., "Diagnoses: Z59.0"), but the NLP logic had been designed to treat those as historical mentions.

Most Veterans who were classified as "Stably Housed" using diagnosis coding but labeled as "Unstably Housed" in the HMIS reference standard clearly documented homelessness in the note but did not have relevant ICD-10 codes associated with the visit. A small number of Veterans who lacked an ICD-10 code for homelessness were actually stably housed but had been incorrectly classified as "Unstably Housed" by both HMIS and the NLP. Differences between HMIS data and EHR documentation could be because the Veteran became stably housed sometime between exiting SSVF and a visit or due to errors in the source HMIS data.

5. Discussion

5.1 Key Findings

This study demonstrates that NLP can be used to detect a Veteran's current housing situation using clinical text. Our method identifies Veterans who are currently unstably housed more accurately than standard diagnosis coding. Additionally, this method is novel in its ability to explicitly identify housing stability in addition to housing instability.

Our document classification system achieved moderate results in classifying documents containing evidence of housing stability ($F1=65.8$). A review of errors and an analysis of the corpus showed that the lexicon was highly variant and ambiguous. Correctly inferring a Veteran's housing situation within a note required detecting linguistic attributes for entities based on contextual information, ignoring templated or irrelevant text, and resolving multiple mentions of housing-related concepts. Our heuristic algorithm outperformed a baseline BOW model and a hybrid NLP/ML model ($F1=50.8/64.1$, respectively). The low performance of these models may be due to the relatively small training set size ($n=412$) which may have been insufficient to train a machine learning model.

The modest inter-annotator agreement observed in our annotated corpus ($Kappa=0.7$) suggests that the task is non-trivial even for domain experts. One factor contributing to this challenge is that the study focused on Veterans who were receiving rapid rehousing assistance during a time of crisis. While there may be a clear difference between Veterans who have been stably or unstably housed for a long period of time, this distinction is much more difficult when the Veteran's housing status is rapidly changing.

Although our system achieved only moderate results at a document level, aggregating to a patient level with *ReHouSED* demonstrated reasonable alignment with reference standard data and showed significantly higher values in the group of stably housed Veterans. Using *ReHouSED* as a classifier for unstable housing achieved an AUROC of 80.2. Comparing our method with standard ICD diagnosis coding showed that our proposed methodology more accurately detects whether a Veteran is currently unstably housed ($F1=44.6$ versus

18.7). In particular, the low precision and specificity of diagnosis coding suggests that many Veterans continue to have diagnosis codes documenting homelessness after they are successfully housed. These findings suggest that this *ReHouSED* could potentially be used as a measurement of housing stability in the absence of reliable structured data. The availability of such a measure is of particular utility because existing measures that capture housing instability are limited to the detection of the absence of housing instability, which does not necessarily imply the presence of stable housing.

Another key finding of our analysis is that Veterans who exited to unstable housing tended to have lower *ReHouSED* upon entry than those who exited to stable housing (median=0.05 versus 0.25). This finding is intriguing because it suggests that SSVF is more successful in housing Veterans with less housing instability at baseline.

There are several potential applications of this method to clinical research. While a Veteran's housing status is recorded in HMIS at the time of discharge from the SSVF program and has been used to assess the short-term impact of different aspects of the SSVF program on housing status [13], there is no systematic approach to capturing a Veteran's housing status in administrative data in the months or even years following SSVF exit. There are many other housing-related interventions in the VA including HUD-VASH and GPD. In evaluation studies, *ReHouSED* could be used to measure Veterans' long-term housing stability following enrollment in these programs. In addition, there are many physical (such as traumatic brain injury or other combat-related injury) or mental health conditions (such as substance use disorder) that may influence a Veteran's ability to maintain stable housing. *ReHouSED* could be used to assess and mitigate the impact of these exposures on Veteran housing status over time. Understanding the long-term impact of interventions can help VA policy-makers target scarce resources toward programs that have a positive impact on Veteran outcomes.

This method could potentially be used outside of VA, where accurate longitudinal assessment of housing stability has presented a persistent methodological challenge in research evaluating the impact of homeless assistance programs. The text processing pipeline described here was designed specifically for processing clinical texts of SSVF participants within VA. While the high-level logic and much of the lexicon for housing-related concepts would be expected to generalize to other datasets and populations, some components of the system may require modification before being applied in new settings. For example, the rules used in the section detection component were designed to fit the structure of VA's EHR and to match the types of notes being authored during a Veteran's participation in SSVF. Similarly, while many terms included in our system's lexicon would apply to a general population, some phrases such as "SSVF housing" or "HUD-VASH voucher" apply only to this specific patient cohort. However, each step of a medspaCy pipeline is highly customizable and rules may be modified or removed to adapt to a new dataset or case definition. Future work should study housing stability in other populations and clinical settings.

An additional finding of this study is the method of aggregating information extracted from multiple clinical notes to describe a patient-level concept. Although the NLP system

which processed each individual document involved complex rules and logic, the method of aggregating results to a patient level to construct the *ReHouSED* score was relatively simple and could be extended to other NLP applications which would benefit from patient-level aggregation.

5.2 Related Work

A number of studies have examined homelessness in clinical text within VA. Gundlapalli et al developed and evaluated a system to extract evidence of homelessness from clinical notes [7]. Their system assigned a binary document classification of whether the Veteran had ever experienced or been at risk of homelessness. In their annotated testing set, they reported precision, recall, and F1 of 94.0, 97.0, and 96.0, respectively, while a review of a larger corpus of 10,000 notes yielded a precision of 70.0. Later work studied a system's precision at extracting mention-level concepts relating to homelessness and associated risk factors, achieving precision scores ranging from 76.0–77.0 [14].

Conway et al developed a system called Moonstone which extracts social risk factors from clinical text [4]. They extracted three document-level classes relevant to housing stability: “Homeless/marginally housed”, “Lives in a facility”, and “Lives at home/not homeless”. They achieved F1-scores of 75.0, 75.0, and 96.0 respectively.

Our work is distinct from prior work in several ways. First, each of the aforementioned studies identified documents showing that individuals had experienced homelessness or instability but did not aim to distinguish between those who are currently homeless and those who have exited homelessness. In contrast, our study focused on a group of Veterans who were in a time of crisis and already known to be experiencing homelessness. This required complex logic to account for recent changes in Veteran housing within clinical narratives. Additionally, none of these studies aggregated information to a patient level, which requires addressing inconsistent information across documents.

5.3 Limitations

Our cohort was limited to Veterans who were known to be experiencing homelessness through active participation in SSVF. The performance of the system was not validated on a general population of patients or on homeless Veterans who were not enrolled in SSVF.

Due to time constraints and the complexity of the task, only document-level classifications were annotated in our sampled corpus of notes. As such, the accuracy of the NLP system for extracting individual entities was not evaluated.

Our NLP document classifier was implemented as a rule-based system. We compared a rule-based algorithm against machine learning models but had a relatively small training set which may have been insufficient for training. Future work should further evaluate the utility of using machine learning to detect housing stability.

Finally, we modeled housing stability as a binary class. Documents containing relevant information were classified as either “Stably Housed” or “Unstably Housed”. However, housing instability is perhaps better thought of as a continuum than a discrete phenomenon.

There is an important distinction between individuals who are at risk of becoming homeless, those staying in temporary emergency housing, or those who are without any shelter. While we initially attempted to make this distinction in document annotation, it was nontrivial to consistently differentiate between these different levels of housing instability. Future work should more closely study different levels of housing instability.

6. Conclusion

Housing stability is an important but complex determinant of health. VA programs such as SSVF support Veterans experiencing homelessness by providing housing assistance, but measurement of long-term outcomes using standard structured data has not previously been possible. In this study, we developed and evaluated a novel patient-level measure of housing stability (*ReHouSED*), which is based in natural language processing and derived by reconciling EHR information across multiple housing-related documents within a window of time. This method aligned effectively with reference standard data and detected Veteran housing status more accurately than standard diagnoses coding. Our findings demonstrate that NLP can be used to measure patient housing stability at specific time points in the absence of reliable structured data. Future work will analyze patterns of housing stability over time and utilize this method to study the long-term outcomes and effectiveness of housing assistance programs.

Acknowledgements

The authors thank Dr. Wendy Chapman, Dr. Adi Gundlapalli, Lee Christenson, and Guy Divita for their insights and feedback on this work.

Funding

This material is the result of work supported with resources and the use of facilities at the George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, Utah. This study was supported with funding from the VA Health Services Research and Development Service [I50HX001240 Center of Innovation - Informatics, Decision-Enhancement and Analytic Sciences (IDEAS) Center and IIR 17-029 (PI: Nelson)]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The views expressed in this paper are those of the authors and do not necessarily represent the position or policy of the U.S. Department of Veterans Affairs or the United States Government.

Audrey L. Jones is supported by an HSR&D Career Development Award (CDA 19-233).

Appendix

Table A.1.

Housing-related terms used in document keyword search.

domiciled	mission	streets
domiciliary	rent	subsidized
evicted	resides	transitional
homelessness	residing	unstably
landlord	shelter	voucher
lease	stably	

Table A.2.

Definitions of document classes for annotation.

Classification	Definitions
Stably Housed	<ul style="list-style-type: none"> • Living in an apartment or home which is paid for by either the patient or VA • Accepted to housing and is preparing to move in • Permanently living with a family member or friend
Unstably Housed	<ul style="list-style-type: none"> • Living in a place not meant for human habitation (the streets, an abandoned building, a vehicle, etc.) • Recently evicted from their current residence • Living in emergency housing or transitional housing • Temporarily staying with a family member or friend • The document does not specify what the patient's exact housing status is, but it is stated that they are facing housing issues or in need of stable housing
Unknown	<ul style="list-style-type: none"> • There is no mention of a patient's housing status • The information in the note is insufficient to make a final judgment

Table A.3.

ICD diagnosis codes used for patient classification comparison.

ICD-9	V60.0	Lack of Housing
	V60.1	Inadequate Housing
	V60.89	Other specified housing or economic circumstances
	V60.9	Unspecified housing or economic circumstance
ICD-10	Z59.0	Homelessness
	Z59.1	Inadequate housing
	Z59.8	Other problems related to housing and economic circumstances
	Z59.9	Problem related to housing and economic circumstances, unspecified

Table A.4.

NLP document classification error analysis.

Error Type	Count	Examples
Mention false negative: A housing-related concepts was either not matched or was ignored	23	<p>“completed hud/vash application” [Evidence of Homelessness]</p> <p>“been accepted for housing” [Evidence of Housing]</p> <p>“housing pursuit” [Evidence of Homelessness]</p> <p>“deliver donated furniture items to Veteran” [Evidence of Housing]</p> <p>“Patient now living with ex-wife” [Evidence of Housing]</p>
Incorrect mention attributes: An attribute such as “is_negated” or “is_hypothetical” was set incorrectly due to a missed modifier or incorrect linking	17	<p>“anxious about moving into his own place” [not recognized as hypothetical]</p> <p>“Veteran requested to be picked up from his apartment” [incorrectly hypothetical]</p> <p>“next week he should be approved for an apartment” [not recognized as hypothetical]</p> <p>“He lived in an apartment but was asked to leave” [not recognized as historical]</p> <p>“Veteran signed his lease using an alias and decided he does not want the apartment” [“signed his lease” recognized as stable housing but is negated later in the sentence]</p>
Potential disagreement with annotator: The classification	16	<p>“Diagnosis: Z59.0” [usually considered historical due to inconsistent coding but sometimes apparently relevant to visit]</p> <p>“Veteran is in housing program through HUDVASH” [marked as</p>

Error Type	Count	Examples
assigned by the annotator could be reasonably disagreed with		stable housing, but may only refer to a treatment program] No explicit information documented in note but marked as Stable/ Unstable
Mention false positive: A phrase of text was matched incorrectly due to lexical ambiguity or incorrect parsing	10	“ Veteran’s place of living ” [extracted as stable housing, later stated to be an abandoned building] “ Domiciliary ” not used consistently in documentation as temporary housing “ his place on the waitlist ” [incorrectly parsed as “his place”]
Other: Poor text formatting, incorrect documentation, or incorrect inference	4	All upper-case text causes incorrect sentence splitting Old information copied and pasted into a note

Table A.5.

The 5 most frequent phrases within the primary entity classes.

Label	Text	Count of Entities	Count of Asserted Entities (% of All Instances)
Evidence of Housing	housing	511	10 (2.0)
	home	213	0 (0.0)
	GPD	122	4 (3.3)
	apartment	108	13 (12.0)
	voucher	91	0 (0.0)
Evidence of Homelessness	homeless	353	230 (65.2)
	HUD-VASH voucher	46	36 (0.78)
	lack of housing	19	0 (0.0)
	sleep in <LOCATION>	20	18 (90.0)
	<RESIDES> <LOCATION>	18	17 (94.4)
Temporary Housing	shelter	66	49 (74.2)
	Xxxx House	51	44 (86.3)
	volunteers of america	43	40 (93.0)
	transitional housing	42	27 (64.4)
	the Domiciliary	17	7 (41.2)
Risk of Homelessness	evicted	16	14 (87.5)
	housing	9	9 (1.0)
	Economic Problem	9	9 (1.0)
	housing needs	8	8 (1.0)
	Unstably housed	8	8 (1.0)
Doubling Up	at <FAMILY> <RESIDENCE>	10	3 (3.0)
	X’s apartment	7	2 (28.6)
	staying with their <FAMILY>	5	5 (1.0)
	couch surfing	3	3 (1.0)

References

- [1]. Hwang SW, Homelessness and health, *CMAJ*. 164 (2001) 229–233. 10.1177/1359105307080576. [PubMed: 11332321]
- [2]. Fazel S, Geddes JR, Kushel M, The health of homeless people in high-income countries: Descriptive epidemiology, health consequences, and clinical and policy recommendations, *Lancet*. 384 (2014) 1529–1540. 10.1016/S0140-6736(14)61132-6. [PubMed: 25390578]
- [3]. Hwang SW, Weaver J, Aubry T, Hoch JS, Hospital costs and length of stay among homeless patients admitted to medical, surgical, and psychiatric services, *Med. Care* 49 (2011) 350–354. 10.1097/MLR.0b013e318206c50d. [PubMed: 21368678]
- [4]. Conway M, Keyhani S, Christensen L, South BR, Vali M, Walter LC, Mowery DL, Abdelrahman S, Chapman WW, Moonstone: A novel natural language processing system for inferring social risk from clinical narratives, *J. Biomed. Semantics*. 10 (2019). 10.1186/s13326-019-0198-0.
- [5]. Gundlapalli AV, Carter ME, Divita G, Shen S, Palmer M, South B, Durgahee BSB, Redd A, Samore M, Extracting Concepts Related to Homelessness from the Free Text of VA Electronic Medical Records, *AMIA Annu. Symp. Proc 2014* (2014) 589–598. [PubMed: 25954364]
- [6]. Redd A, Carter M, Divita G, Shen S, Palmer M, Samore M, Gundlapalli AV, Detecting earlier indicators of homelessness in the free text of medical records, *Stud. Health Technol. Inform 202* (2014) 153–156. 10.3233/978-1-61499-423-7-153. [PubMed: 25000039]
- [7]. Gundlapalli AV, Carter ME, Palmer M, Ginter T, Redd A, Pickard S, Shen S, South B, Divita G, Duvall S, Nguyen TM, D'Avolio LW, Samore M, Using natural language processing on the free text of clinical documents to screen for evidence of homelessness among US veterans., *AMIA Annu. Symp. Proc 2013* (2013) 537–546. [PubMed: 24551356]
- [8]. Jones B, Gundlapalli AV, Jones JP, Brown SM, Dean NC, Admission decisions and outcomes of community-acquired pneumonia in the homeless population: A review of 172 patients in an urban setting, *Am. J. Public Health*. 103 (2013). 10.2105/AJPH.2013.301342.
- [9]. Gelberg L, Andersen RM, Leake BD, The Behavioral Model for Vulnerable Populations: application to medical care use and outcomes for homeless people., *Health Serv. Res 34* (2000) 1273–302. <http://www.ncbi.nlm.nih.gov/pubmed/10654830> (accessed March 24, 2021). [PubMed: 10654830]
- [10]. Fargo J, Metraux S, Byrne T, Montgomery AE, Jones H, Culhane D, Munley E, Sheldon G, Kane V, Prevalence and risk of homelessness among US veterans, *Prev. Chronic Dis 9* (2012). 10.5888/pcd9.110112.
- [11]. Evans WN, Kroeger S, Palmer C, Pohl E, Housing and urban development-veterans Affairs supportive housing vouchers and veterans' homelessness, 2007–2017, *Am. J. Public Health*. 109 (2019) 1440–1445. 10.2105/AJPH.2019.305231. [PubMed: 31415190]
- [12]. Nelson RE, Suo Y, Pettey W, Vanneman M, Montgomery AE, Byrne T, Fargo JD, Gundlapalli AV, Costs Associated with Health Care Services Accessed through VA and in the Community through Medicare for Veterans Experiencing Homelessness, *Health Serv. Res 53* (2018) 5352–5374. 10.1111/1475-6773.13054. [PubMed: 30246368]
- [13]. Nelson RE, Byrne TH, Suo Y, Cook J, Pettey W, Gundlapalli AV, Greene T, Gelberg L, Kertesz SG, Tsai J, Montgomery AE, Association of Temporary Financial Assistance With Housing Stability Among US Veterans in the Supportive Services for Veteran Families Program, *JAMA Netw. Open 4* (2021) e2037047. 10.1001/jamanetworkopen.2020.37047.
- [14]. Peterson R, Gundlapalli AV, Metraux S, Carter ME, Palmer M, Redd A, Samore MH, Fargo JD, Smalheiser NR, Identifying homelessness among veterans using VA administrative data: Opportunities to expand detection criteria, *PLoS One*. 10 (2015). 10.1371/journal.pone.0132664.
- [15]. Cohen J, A Coefficient of Agreement for Nominal Scales, *Educ. Psychol. Meas 20* (1960) 37–46. 10.1177/001316446002000104.
- [16]. Eyre H, Chapman AB, Peterson KS, Shi J, Alba PR, Jones MM, Box TL, DuVall SL, V Patterson O, Launching into clinical space with medspaCy: a new clinical text processing toolkit in Python, *AMIA Annu. Symp. Proc 2021* (in Press. (n.d.)). <http://arxiv.org/abs/2106.07799>.

- [17]. Chapman A, Peterson K, Turano A, Box T, Wallace K, Jones M, A Natural Language Processing System for National COVID-19 Surveillance in the US Department of Veterans Affairs, in: Proc. 1st Work. NLP COVID-19 ACL 2020, 2020.
- [18]. Harkema H, Dowling JN, Thornblade T, Chapman WW, ConText: an algorithm for determining negation, experiencer, and temporal status from clinical reports., *J. Biomed. Inform* 42 (2009) 839–51. 10.1016/j.jbi.2009.05.002. [PubMed: 19435614]
- [19]. Chen T, Guestrin C, XGBoost: A scalable tree boosting system, in: Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., Association for Computing Machinery, 2016: pp. 785–794. 10.1145/2939672.2939785.

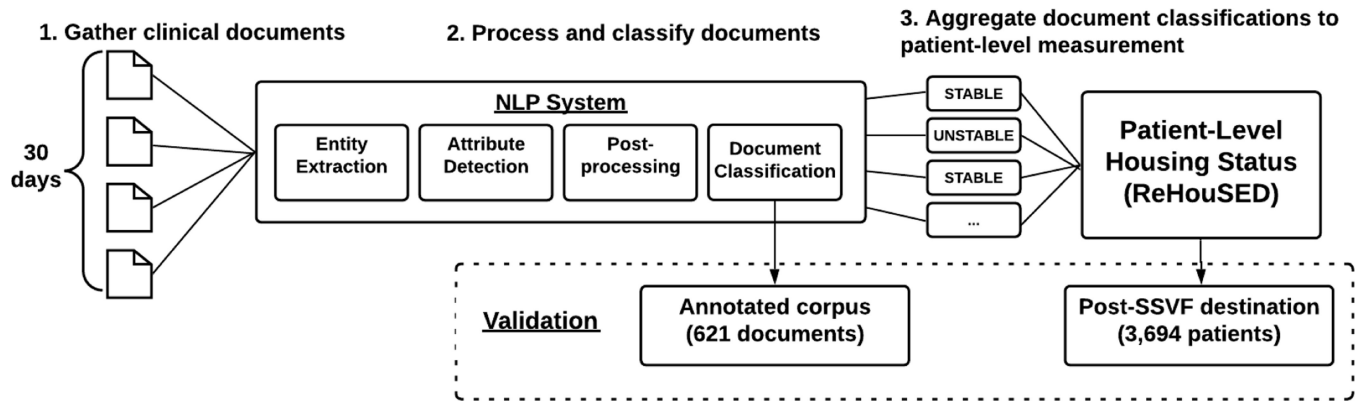


Figure 1.

The steps for processing notes and calculating patient-level Relative Housing Stability in Electronic Documentation (*ReHouSED*) scores. Sources for reference standard data used for system evaluation are shown on the bottom.

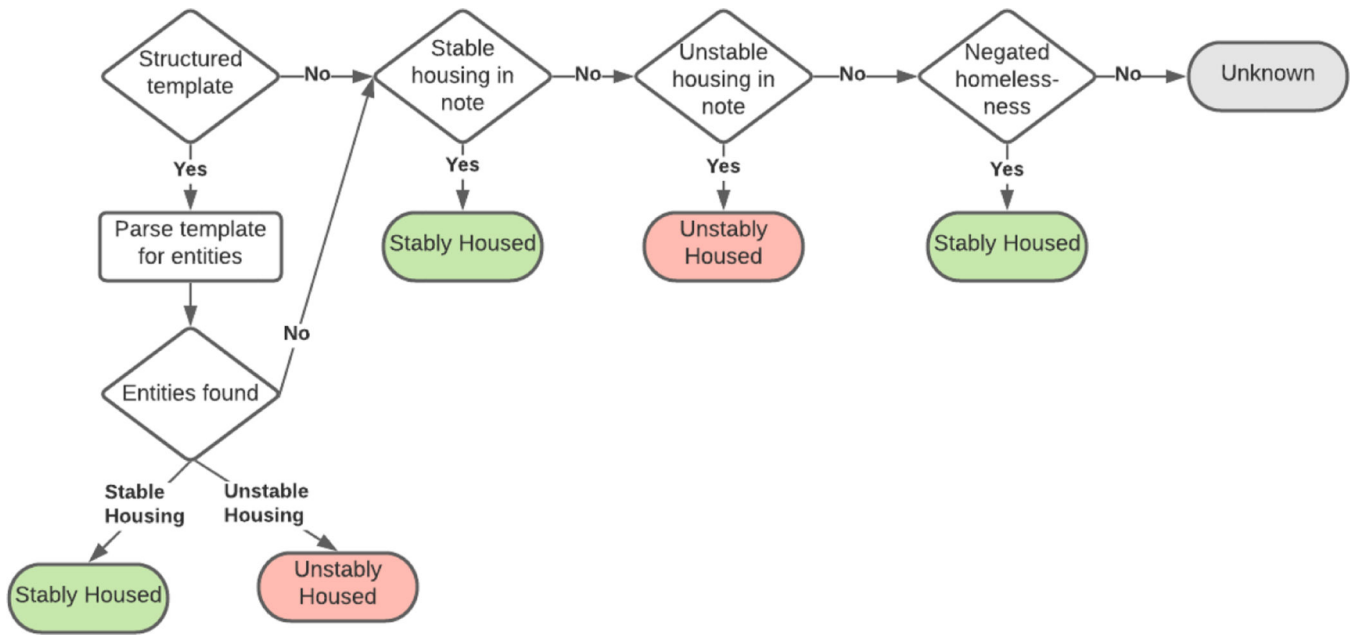


Figure 2. Natural language processing (NLP) document classification logic.

VA Author Manuscript

VA Author Manuscript

VA Author Manuscript

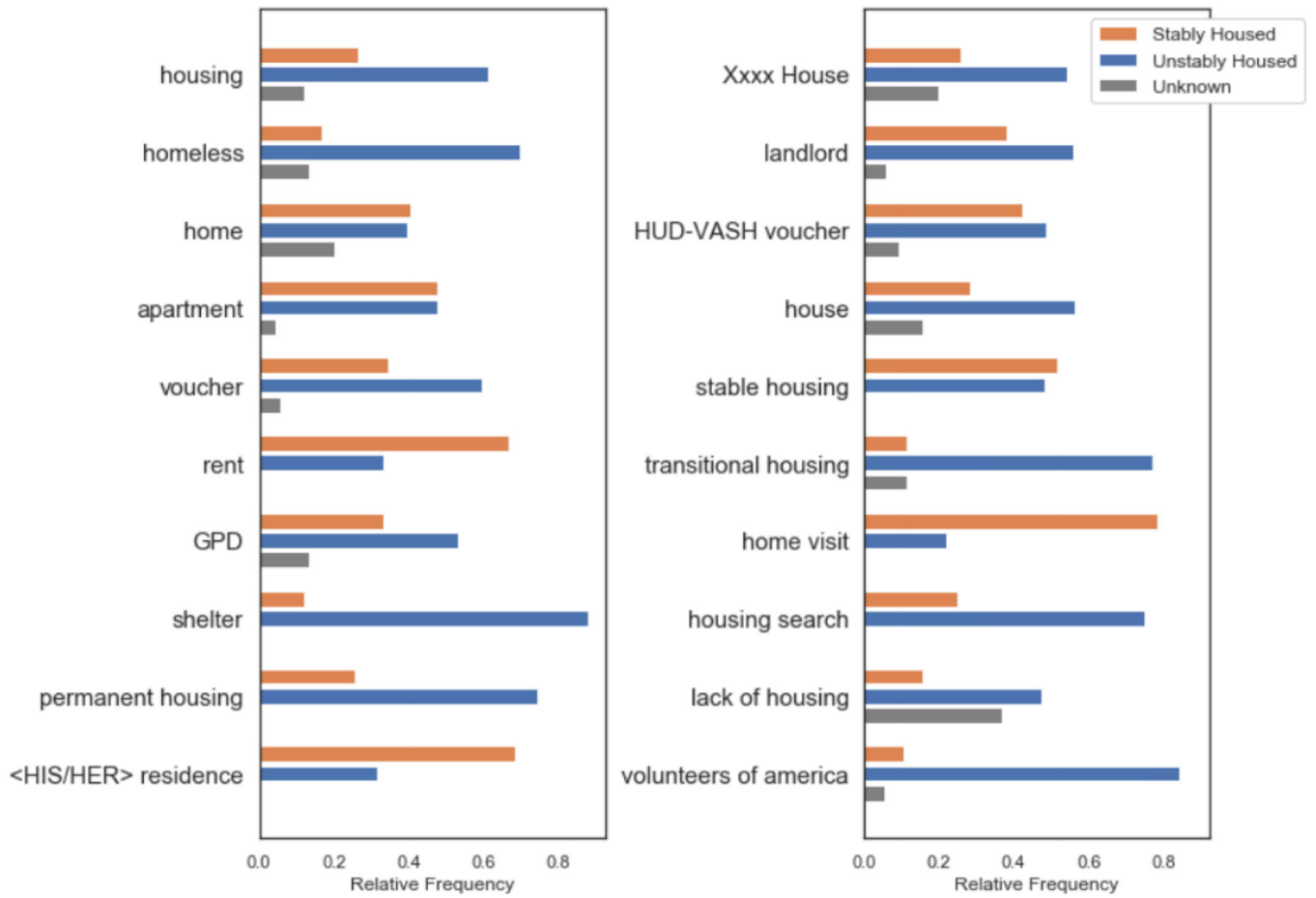


Figure 3.
The relative frequency of phrases extracted from the corpus by annotated document class.

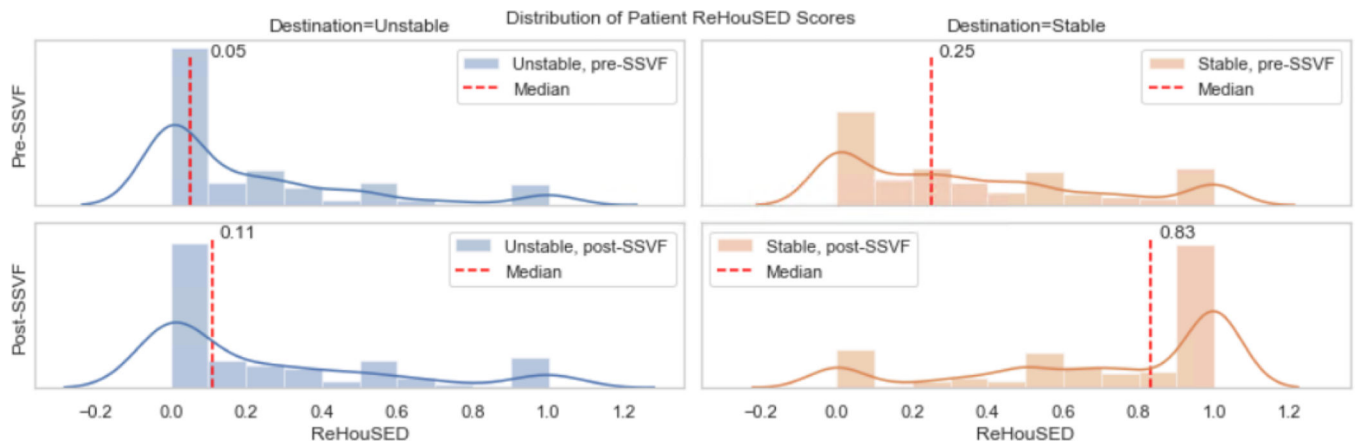


Figure 4.

Histograms showing the distribution of *ReHouSED* values by time period (pre- and post-SSVF) and housing destination (Unstable and Stable). SSVF=Supportive Services for Veteran Families; *ReHouSED*=Relative Housing Stability in Electronic Documentation.

Table 1.

Counts of Veterans and treatment episodes in Supportive Services for Veteran Families (SSVF).

Total number of Veterans	86,186
Total number of SSVF episodes	95,022
Destination=Stably Housed	78,264 (82.3%)
Destination=Unstably Housed	16,758 (17.7%)

Table 2.

Examples of the five entity classes extracted by our system. Patterns used for entity extraction are shown next to sentences containing spans of text matched by the rule (in **bold**).

Entity class name	Pattern	Examples
Evidence of Stable Housing	"<PRON> <ADJ> apartment" "<PAYS> rent"	"The veteran is doing well in her new apartment. " "He is current on the rent. "
Evidence of Homelessness	"admitted from the streets" "literally homeless"	"Veteran admitted from the streets. " "The patient is currently literally homeless. "
Temporary Housing	"the Mission" "shelter"	"Spent last night at the Mission. " "Got a bed at a shelter downtown."
Doubling Up	"<STAY> <FAMILY>" "<STAY> <FRIEND>"	"His mother let him stay with her. " "She crashed at a friend's house. "
Risk of Homelessness	"<NOT> pay rent" "eviction notice"	" Cannot pay the upcoming rent. " "Got an eviction notice. "

Table 3.

Examples of entities (in **bold**) and attributes (in *italics*). The entity class was only considered to be asserted if it was not assigned any of the first four attributes.

Attribute name	Example	Entity class
is_negated	"The patient is <i>not currently</i> homeless ."	Homelessness
	"She does <i>not have a place to stay</i> tonight."	Stable Housing
is_hypothetical	"He <i>would like</i> to have his own apartment ."	Stable Housing
	"Concerns: <i>worried about becoming</i> homeless "	Homelessness
is_historical	" <i>PMH:</i> Homeless single person"	Homelessness
	" <i>Previously</i> had his own apartment ."	Stable Housing
is_ignored	"Here to <i>discuss</i> hishousing situation."	Stable Housing
	"Is the patient homeless ? [] Y [] N"	Homelessness
is_asserted	"She is homeless ."	Homelessness
	"She <i>lives in an</i> apartment ."	Stable Housing

Table 4.

NLP document classification performance on a blind test set.

	Precision	Recall	F1	Number of documents
Stably Housed	64.0	62.7	63.4	51
Unstably Housed	78.3	64.9	70.9	111
Unknown	53.7	76.6	63.2	47
Macro Avg	65.3	68.1	65.8	209

Table 5.

Document classification baseline comparison.

	Precision	Recall	F1
NLP	65.3	68.1	65.8
BOW	51.6	50.6	50.8
NLP + ML	64.1	64.1	64.1

Table 6.

Steps to assemble sample of 3,694 Veterans and 35,452 documents to evaluate Relative Housing Stability in Electronic Documentation (*ReHouSED*) score.

	Number of Veterans	Number of Documents
1. Initial sample of Veterans	10,328	--
2. Retrieve documents from pre- and post-SSVF time intervals	7,131	62,028
3. Exclude documents classified by NLP as "Unknown" and exclude Veterans without at least one document in each time interval	3,694	35,452

Table 7.

Median *ReHouSED* values before and after participation in SSVF, by housing status recorded at the time of SSVF program exit. IQR=Interquartile Range; HMIS=Homeless Management Information. Systems; SSVF=Supportive Services for Veteran Families; *ReHouSED*=Relative Housing Stability in Electronic Documentation.

	Housing Status Recorded in HMIS				P-value
	Unstable (N = 415)		Stable (N = 3,279)		
	Median ReHouSED	IQR	Median ReHouSED	IQR	
Pre-SSVF	0.05	(0.0–0.33)	0.25	(0.0–0.54)	<0.001
Post-SSVF	0.11	(0.0–0.4)	0.83	(0.5–0.83)	<0.001

Table 8.

Binary classification of Veteran housing status using the NLP-derived Relative Housing Stability in Electronic Documentation (*ReHouSED*) measure versus ICD diagnosis codes.

	Precision	Recall	F1	NPV	Specificity	AUROC
ReHouSED	32.1	72.9	44.6	95.9	80.3	80.2
Diagnosis codes	10.8	69.4	18.7	87.4	27.1	--