2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Leveraging Unsupervised Machine Learning to Discover Patterns in Linguistic Health Summaries for Eldercare

Pallavi Gupta[1], Omar Ibrahim[2], Marjorie Skubic,[2] *Senior Member, IEEE,* and Grant J. Scott[1], *Senior Member, IEEE*

Abstract— **The Center for Eldercare and Rehabilitation Technology, at University of Missouri, has researched the use of smart, unobtrusive sensors for older adult residents' health monitoring and alerting in aging-in-place communities for many years. Sensors placed in the apartments of older adult residents generate a deluge of daily data that is automatically aggregated, analyzed, and summarized to aid in health awareness, clinical care, and research for healthy aging. When anomalies or concerning trends are detected within the data, the sensor information is converted into linguistic health messages using fuzzy computational techniques, so as to make it understandable to the clinicians. Sensor data are analyzed at the individual level, therefore, through this study we aim to discover various combinations of patterns of anomalies happening together and recurrently in the older adult's population using these text summaries. Leveraging various computational text data processing techniques, we are able to extract relevant analytical features from the health messages. These features are transformed into a transactional encoding, then processed with frequent pattern mining techniques for association rule discovery. At individual level analysis, resident ID 3027 was considered as an exemplar to describe the analysis. Seven combinations of anomalies/rules/associations were discovered in this resident, out of which rule group three showed an increased recurrence during the COVID lockdown of facility. At the population level, a total of 38 associations were discovered that highlight the health patterns, and we continue to explore the health conditions associated with them. Ultimately, our goal is to correlate the combinations of anomalies with certain health conditions, which can then be leveraged for predictive analytics and preventative care. This will improve the current clinical care systems for older adult residents in smart sensor, aging-in-place communities.**

I. INTRODUCTION

The Center for Eldercare and Rehabilitation Technology (CERT) was founded with the aim of proactive healthcare and the vision to help the senior adults living alone lead a healthier and more independent life. CERT comprises an interdisciplinary group of people consisting of nurses, doctors, and engineers working together in a collaborative effort to develop new health care technologies for eldercare [1]. The research done at CERT is driven by actual clinical needs and their evaluations in the realistic settings at aging-in-place facilities.

Americare facilities, such as *Tiger Place*, are a key example [2] and the source of data used within this study*. The older adult residents live independently under the monitoring of non-wearable, unobtrusive smart sensors [3], [4]. Smart home technologies connected with the internet of things (IOT) facilitate advanced older adult residents health monitoring and connects the residents to the outer world, for a better outreach, especially in case of emergency, without overt and unnecessary intervention [5]. Continuous health monitoring is maintained using bed sensors, motion sensors, and depth sensors, as well as advanced algorithms that derive physiological readings [6] from the sensor suite. Residents, through their informed consent, form a holistic, analytics-driven, smart care community.

Diverse, multi-modal sensor data from facilities such as this represent significant big data challenges for archiving, access, and analytics. The system data is composed of raw data layers, algorithmically generated features, as well as computationally produced analytical information layers. Herein, we seek to explore the development of advanced analytical methods over the top of existing analytical layers. Using smart sensors data to discover health anomalies is one layer of analytics, which has been accomplished in earlier research [7]; and then, finding combinations of patterns of those health anomalies occurring together and recurrently is a secondary (layered) analytical challenge. This paper specifically tackles the later challenge, discovering trends of combinations of patterns of anomalies recurring over time in older adult residents using association rule mining (ARM).

ARM is an unsupervised machine learning, data mining approach that helps in discovering patterns and correlations in data [8]. Other research efforts have leveraged ARM for feature identifications [9] or linguistic memorization of data sets using association rules [10]. However, to the best of our knowledge, ARM has not previously been used as a health anomalies trend discovery tool in the context of linguistic health summaries analyzing both co-occurrences and recurrence patterns. We have leveraged ARM in a novel approach by transforming health messages into a transactional data set, representing a variety of residents, analytical health information, and trends.

This paper is organized as follows. We introduce related research elements from CERT in Sect. II, followed by a detailed discussion of our methodologies in Sect. III. Section IV discusses our experimental results and findings. Then, we provide concluding remarks in Sect. V.

II. RELATED WORK

As noted above, endeavours and knowledge of people from various domains converge at Americare facilities, an analytics-driven smart care community for older adults. This

[1]Gupta and Scott are with Institute for Data Science and Informatics, University of Missouri, Columbia, MO 65211, USA (GrantScott@missouri.edu)

[2]Ibrahim and Skubic are with the Center for Eldercare and Rehabilitation Technology, University of Missouri, Columbia, MO 65211, USA

*Data collection is done with the University of Missouri IRB approval.

includes engineering of robust, unobtrusive sensing systems, as well a care-oriented computational systems. Linked research efforts include many important topics for eldercare, such as fall detection, health monitoring, and clinical care. Herein, we focus specifically on a computational linguistic summarization of health data as a layer of information for secondary analytical processing.

A. Smart Sensing Eldercare Environment

There are variety of sensors being used at Americare facilities to derive physiological readings, including bed sensors, motion sensors (in living room, bathroom, kitchen etc.), and depth sensors [11]. Bed sensors, specifically a ballistocardiogram (BCG), extract high-frequency readings that are processed to generate resting heart rate [12], [13], [14], [15], respiration rate [13], [16], time in bed and bed restlessness [13]. Motion sensors help to capture overall activity, such as the time in spent in a bathroom [17]. Depth sensors feed advanced algorithms that capture kinetic motion measurements, such as stride time, stride length and walking speed [18], [19].

Ultimately there are nine major features which are being constantly monitored to characterize the health and well-being of older adult residents. These nine features include pulse rate, respiration rate, walking speed, bed restlessness, time in bed, stride time, stride length, overall activity, time spent in bathroom. These behavioural features may not seem very critical; but, whenever health conditions are encountered by a person, deviations in these behavioural routine activities of day-to-day life tend to be observable, especially in senior persons [20]. For example, more visits to bathroom means more time spent in bathroom, which may, for instance, represent the possibility of urinary tract infection (UTI). Hence, by analyzing these behavioural features, the health conditions of the older adults can be studied.

B. Linguistic Summaries and Alerting

Each day, data from the sensors are analyzed automatically. If anything in these nine features goes above or below the normal limits, it is detected as an anomaly and textual health summary messages (alerts) are generated. Needless to say, considering advance age residents, there are numerous health, life, and other circumstances that can lead to the generation of health messages. Over the last few years, this has resulted in a large corpus of textual summary alerts.

The text summaries are generated using fuzzy computational linguistics [7] and, in their own right, is an analytical layer. The anomalies found in residents behaviour using sensors data are converted to text, using fuzzy methodologies to provide clinicians with a more informative, accurate and concise reports, whenever needed. However, this only solved the short-term purpose of understanding the anomalies.

Considering the bigger picture, by examining the corpus of linguistic summaries, a higher-level analytical layer can be generated that provides insights into trends and conditions beyond what is apparent from the intermediate features that the summaries are derived from. Hence, we seek to discover trends and anomaly patterns within the summaries that will facilitate more proactive care for older adult residents.

TABLE I
EXAMPLE LINGUISTIC HEALTH SUMMARIES

Type	Message
1	In the past two weeks, there were many days with low time spent in bathroom and low stride length, low walking speed and high stride time, low overall activity and low respiration rate.
2	Night-time bed restlessness, night-time time in bed have been increasing for the past 5 days.
3	Yesterday, time in bed was significantly higher than usual.

To facilitate analysis of linguistic health summaries, we identified and categorized them into three types of textual health messages.

1) *Type 1*: Generated via a sliding window that examines the past two weeks, where significantly anomalous trends are identified within the window.
2) *Type 2*: Generated when there is a continuous accumulation of a certain anomaly for more than five days, but less than two weeks. These are to keep track of the anomalies which subside before two weeks.
3) *Type 3*: Generated on the regular daily basis when any anomaly is encountered such that normal, expected variances are exceeded, such as measurements outside of two standard deviations of a historic mean value. These measures are prone to false alarms related to spurious activities or residents having guests.

Table I provides examples of each type of health message. The subsequent aspects of this current work focuses on *Type 1* health messages.

III. ARM FOR LINGUISTIC HEALTH SUMMARIES

When it comes to analysing the textual information, it is intuitive to lean towards natural language processing (NLP) techniques for feature extraction. We initially investigated the use of a variety of NLP techniques for feature extraction, but ultimately concluded that the health message structure and its relation to source features were better suited for an alternative approach. An alternative perspective is to consider the health summaries as time series data. That is, *Type 1* and *2* are, in essence, a moving average type aggregation of physiological health features over a two week and five day period, respectively. However, the messages are generated at irregular patterns, driven by data events, instead of regular sampling intervals. Even *Type 3* health messages are irregular, despite their potential for daily generation with single event associations. Basically, traditional time-series analytics methods do not well model the linguistic health messages. Therefore, we examined the corpus of summaries with a transactional encoding, detailed below, analyzing both the individual level and population level to discover associations.

A. Individual Level Analysis

As noted, *Type 1* health messages represent physiological health features that are tracking a two week trend and

TABLE II

LINGUISTIC SUMMARY FEATURES

high/low	time in bed	high/low	bed restlessness
high/low	overall activity	high/low	pulse rate
high/low	respiration rate	high/low	stride length
high/low	stride time	high/low	walking speed
high/low	time spent in bathroom		

therefore significant. This significance is why our initial work focuses on mining information on these messages. That is, *Type 1* messages are most likely to capture acute health conditions that do not subside within two weeks and, therefore, may evolve into chronic conditions. This long view analytic also rules out the likelihood of a false alarm, i.e., non-significant event.

In short, *Type 1* textual summaries are most reliable to draw conclusions related to health conditions in older adults residents. Hence, we extracted all *Type 1* textual summaries from the system. This included 1185 total messages, spanning 21 months, for 87 residents. It should be noted that there is little difference expected in the processing of *Type 2* and *3* relative to *Type 1*. It should also be noted that there is no difference in the feature extraction for the purpose of individual level versus population level analytics.

1) Linguistic Summary Feature Extraction: Table II shows the features and modifiers used for *Type 1* health messages. Modifiers are straight forward, *high* or *low*, describing the two week trend relative to prior normal levels. The underlying data analytic is the other component of the feature: “time in bed”, “bed restlessness”, “overall activity”, “pulse rate”, “respiration rate”, “stride length”, “stride time”, “walking speed”, and “time spent in bathroom”. In this case the *Type 1* textual summaries were extracted for each resident of interest and then features were extracted. The historical linguistic health messages archive, i.e., corpus, is organized as a (date, time, resident id, message) collection. For each message, using the combination of modifier and analytic can encode up to 18 features as a transaction for that day. For example, a message may include the textual information “The average walking speed over the last week was significantly lower than usual.” This results in the feature extraction for *low walking speed*. If that same user, on the same day has a health message “The average time in bed over the last two weeks is higher than usual”, then *high time in bed* will extracted. Once features were extracted, transactional encoding is applied to the features with the *date of occurrence* of the anomalies as the transaction index or *item id*. Figure 1 shows how the health features from Table II become binary transactions. Recall, this analysis is performed on a per-individual (resident) level.

2) Applying Pattern Mining using Association Rules: Apriori algorithm is a popular technique for mining frequent item sets, having relevant associations in various transactions. *Item sets* refers to the combinations, or subsets, of items from a super-set of items of interest that occur or reoccur together in various transactions. A *Transaction* refers to the instances of different item-sets occurring or reoccurring in different

Extracted Features

Transactions

	high pulse rate	low pulse rate	high time in bed	low time in bed	high respiration rate	low respiration rate	high stride time	low stride time	high stride length	low stride length	high walking speed	low walking speed	low bed restlessness	high bed restlessness	low overall activity	high overall activity	high time spent in bathroom	low time spent in bathroom
T1	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
T2	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
T3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
T5	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1
T6	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0

Fig. 1. Example of transactional encoding for *Type 1* linguistic health summaries.

patterns or combinations, having unique identifiers. In our scenario, the unique identifiers are the combination of date and resident identifier.

Apriori algorithm is a multi-staged algorithm, wherein the aim is to start with single items in the set and then keep the chain of combinations of items growing, iteratively, depending upon if the combination is found to have support above a threshold. The chain keeps growing until the last longest chain of combination of items is found with the necessary support. Since the process is iterative, each iteration generates some rules and there would be instances in results where the shorter chains of combinations of items are subsets of longer chains; also there may be some repetitions with different permutations of the same items. Therefore, pruning of the discovered item chains is necessary to remove redundancy.

The resulting combination of items found after applying Apriori algorithm are called *Rules*. The relevance of associations in rules is decided on the basis of metrics like *lift*, *support*, and *confidence*. Associations in rules are usually represented in the form of antecedents and consequent. This can be represented as:

$$A(antecedent(s)) \rightarrow B(Consequent) \tag{1}$$

This means, in a rule $A \rightarrow B$, occurrence of A antecedent(s) would result in occurrence of B consequent. There can be many such different $A \rightarrow B$ combinations/rules with each one having associated measures of lift, support and confidence. Lift defines the correlation of A antecedent(s) with B consequent in rule $A \rightarrow B$. For *lift* > 1, A antecedents and B consequent are positively correlated; but for *lift* < 1, the correlation is negative.

Let’s consider $n(A)$ is the count of transactions with item set A, $n(B)$ is the count of transactions with item set B, $n(A \cap B)$ is the count of transactions containing A and B. Then *lift* for potential rule $(A \rightarrow B)$ is calculated as

$$Lift = \frac{n(A \cap B)}{n(A) * n(B)}. \tag{2}$$

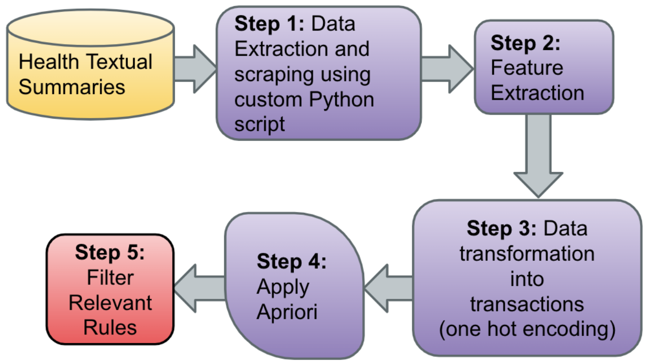

Fig. 2. ARM Pipeline for individual level analysis: The *Type 1* health message corpus is encoded into transactions and association rules are mined on an individual level.

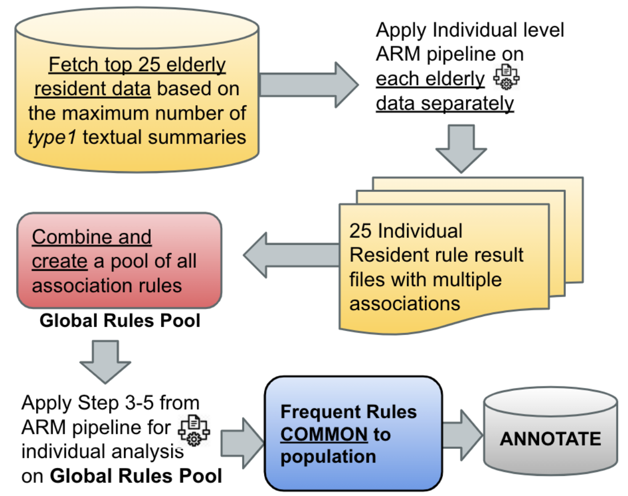

Fig. 3. ARM Pipeline for linguistic health messages: The health message corpus is encoded into transactions and association rules are mined from per-individual perspective; followed by agglomeration for population level analysis.

Recall, that both A and B are item sets, and therefore $\cap$ represents co-occurrence, or intersection, of the two sets. *Support* defines the frequency of recurrence of the respective rule $A \rightarrow B$. It measures how frequent or common a rule is in a corpus of transactions.

$$Support = \frac{n(A \cap B)}{N}, \tag{3}$$

where N is total number of transactions in a dataset. *Confidence* defines the conditional probability, signifying, if A antecedent(s) occurred, how many times B consequent was also within the transaction. There may be transactions when A was present, but B was not, and thus, this combination would not be found despite the presence of A antecedent(s). Confidence is defined as

$$Confidence = \frac{n(A \cap B)}{n(A)}. \tag{4}$$

In this research, the 18 features detailed in Table II form the items super-set containing all the possible health items. The textual summaries generated on a date for a particular resident represented a transaction with date as the transaction identifier; and then the features are one hot encoded to be the items within the transaction. In this setting, the items are health features or anomalies attributed to a person on a given date. Once the linguistic summary health messages are encoded, the corpus has been transformed into transactions for association rule mining.

Once rules were obtained, they were filtered for the unique rules, having *lift* $>= 1$ and number of *antecedents* plus *consequent* $>= 3$. Our interest is in the rules where *antecedents* and *consequent* that show a positive relation, therefore *lift* $>= 1$ is used as a filter. We did not filter the rules based on support because maximum *support* anticipated for any rule was 10-12%; the reason being that, there could be multiple health conditions occurring in an older adult, and each health condition is expected to have a combination of anomalies, i.e, rules, associated with it.

B. Population Level Analysis

For our initial population level analysis, we desired a corpus of messages that had high transaction density based on the number of *Type 1* textual summaries they had generated. That is, the residents with the most rich linguistic health information. One of the primary goals is to discover interesting, yet irregular, temporal patterns within the population, so we constrained the population analysis to the data of the top-25 *Type 1* message generating residents. Residents having a lower number of textual summaries would skew the data more towards rare outliers, thereby detracting from discovering critical trends.

Once the *top-25* linguistic health corpus transactions were extracted, a similar pipeline as the individual level analysis is performed for each resident; except that the rules were filtered for *lift*$>= 1.5$ instead of 1 as threshold. The filtered rules are then combined to form a global rules pool. A second pass of association rule discovery is then applied on these global rules to get the frequent common rules of the population with filtering *lift*$>= 1$. This leads to the discovery of some rules that are highly specific to the individuals, depending upon their medical history, but also key rules common to the *top-25* residents. This population level analysis is able to discover frequently occurring patterns in the population, and, when considered with temporal context can be enlightening towards outside factors affecting an older adult population within a smart facility.

IV. Experimental Findings

As mentioned earlier we were interested in discovering trends at individual and population level. Since we were looking only for positively correlated antecedents and consequent we chose the lift value as the metric for filtering association rules. It is expected not to have very high support for each rule, since each combination of anomalies likely represents only one health condition in older adults and it is expected to have multiple such health conditions reoccurring over time in older adult residents.

Rule	Rule Group
high pulse rate', 'low bed restlessness', 'low time in bed'	7
high pulse rate', 'high respiration rate', 'high stride time'	6
'high time in bed', 'high time spent in bathroom', 'low overall activity'	5
'high time in bed', 'low overall activity', 'low stride length'	4
'high pulse rate', 'low bed restlessness', 'low time spent in bathroom'	3
high respiration rate', 'low bed restlessness', 'low overall activity'	2
high pulse rate', 'high respiration rate', 'low stride length'	1

Table of 7 Rules found in 3027

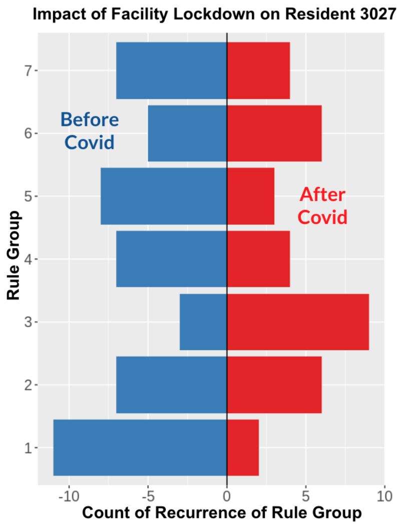

Fig. 4. Top 7 rules found in Resident ID 3027 and their recurrence count; Left: Table containing top 7 rules; Right: Plot for respective rule recurrence count on various dates (before and after COVID-19 facility lockdown).

A. Individual Level Analysis

While we examined multiple residents' data, here we focus on one particular resident, **3027** (a 96 year old female), for illustrative individual results. In total 142 rules were discovered for resident identifier 3027, and this includes repetitions, i.e., alternative rule permutations, as well. Hence, rules were filtered to keep the unique rules, having *lift* > 1, and antecedents plus consequent count > 2. This filtering reduces our investigation of resident 3027 to seven key rules.

Given that the discovered rules illustrate combinations of patterns recurring over time, we explored the effects of a COVID-19 facility lockdown on the health messages. Figure 4 displays these seven rules and their recurrence counts before and after COVID facility lockdown. The health messages for resident 3027 cover the time from May 04, 2019 to August 31, 2020; and on March 13, 2020 the facility entered lock-down protocols to protect the residents' health. In Fig. 4, the occurrences of each rule group before (blue) and after (red) lockdown are visualized. It can be clearly seen that *Rule3*, including anomalies 'high pulse rate', 'low bed restlessness', and 'low time spent in bathroom' become more prevalent with the onset of COVID-19 lockdown. Conversely, we see that *Rule1* had an opposite occurrence shift based on the COVID-19 lockdown, where the number of generated health messages decreased significantly. Qualitatively, these two results are very intuitive. During the facility lock-down, *Rule1* which is indicative of physical activity (high pulse rate and respiration), see a decrease; whereas, *Rule3* is more indicative of an in-room confinement situation. Figure 5 shows the dates of recurrence of *Rule3* in resident ID 3027 before and after the facility lock-down, with blue bordered representing dates before lockdown, and red bordered represents dates after the lockdown.

Fig. 5. Dates of recurrence of Rule 3 in resident ID 3027 before (bounded blue) and after facility COVID-19 lockdown (bounded red); *Rule 3*:'high pulse rate','low bed restlessness','low time spent in bathroom'.

Rule: 'high stride time', 'low stride length', 'low walking speed'.

Resident ID	Recurrence Dates for Rule
3132	2019-09-28
	2019-12-07
6018	2019-06-01
	2019-06-10
6422	2019-11-10
	2020-01-01
3038	2019-08-14
	2019-09-18

Acute Condition Cases!

Resident ID	Recurrence Dates for Rule
3054	2019-09-06
	2019-12-12
	2020-03-26
	2020-03-28
	2020-04-02
	2020-04-03
3027	2019-09-17
	2020-05-01
	2020-05-02
	2020-06-06
	2020-06-07
3109	2019-05-24
	2020-07-04
	2020-08-24
	2020-09-01

Chronic Condition Cases!

Fig. 6. Dates of recurrence of an example rule in the population.

B. Population Level Analysis

After filtering the global rules using the same constraints as described above for individual rules, 38 population-level rules are extracted. However, at the population level, splitting *rule* occurrences into *before* and *after* the COVID-19 facility lockdown is ill-suited, as individuals will have different personal reactions to the circumstances. Since individuals behave differently towards pandemic protocols, if the frequency of certain *rules* or conditions show an increase in one individual, it is possible that another individual will behave completely opposite. In Fig. 6, we examine a single *rule*: 'high stride time', 'low stride length', and 'low walking speed'. There was recurrence of this *rule* at the population level in various residents (seven shown in the figure with dates of occurrence).

Exploring the *rule* patterns in Fig. 6, we identified two characteristic sets of residents. In a first group, residents 3132, 6018, 6422, 3038, we identified transient, acute instances of the condition. However, the other group, residents 3027, 3054, 3109, have more occurrences over longer time spans in the resident history; these are identified as chronic in nature. This clearly shows how individuals differ from one another, and in the eldercare setting, generalizing findings at the population may not be as effective as individualized analysis. Hence, individual-level precision analysis appears to have more promise based on our observations to date.

V. Conclusion

At Americare facilities, such as *Tiger Place*, constant monitoring is performed by automatically aggregating, analysing, and summarising the deluge of daily sensor data for residents; and if any statistical anomaly is encountered, linguistic summaries are generated describing the condition as a health message. We have presented an initiative to discover patterns of combinations of such anomalies within the health messages for particular residents (individualized) and across all residents (global). We extracted 18 transactional features, each representing an anomaly, from the linguistic summaries. ARM was then applied to these transactions to find the rules or anomalies that occurred together, at the individual level as well as at the population level.

To elucidate the individual level analysis, an exemplar resident case was discussed and shown that seven key anomaly patterns were prevalent from the past approximately two years of linguistic summary data. Additionally, we examined the impact of COVID-19 facility lockdown, examining how the occurrences of these seven rules changed due to the lockdown.

At the population level, 38 rules were found common to the residents whose data was used for this analysis. When examining population level rules, recurrence patterns of rules associated to a particular resident allowed us to see indications towards chronic and acute health conditions. Numerous individuals sharing similar combination of anomalies in either the chronic or acute recurrence can further be analyzed to determine what, if any, underlying health conditions may lead to these health messages.

Overall, we were able to discover recurring patterns of combinations of anomalies using *Type 1* linguistics summaries data from the past. However, annotating or associating these combination of anomalies with specific health conditions, like urinary tract infection or heart disease is our more lofty goal, as it would be even more beneficial for early prediction of diseases in future for eldercare. Hence future work would involve annotating these combinations of anomalies by discovering linkages within the electronic health records. Additionally, integrating the *Type 2* and *Type 3* linguistic health messages into this endeavour for predictive analytic and preventative care is a critical future step.

Acknowledgement

This work was supported in part by the National Institute of Nursing Research of the National Institutes of Health under award number R01NR016423. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

[1] M. J. Rantz, K. D. Marek, M. Aud, H. W. Tyrer, M. Skubic, G. Demiris, and A. Hussam, "A technology and nursing collaboration to help older adults age in place," *Nursing Outlook*, vol. 53, no. 1, pp. 40–45, 2005.

[2] M. Fergenson, "Tigerplace: An innovative 'aging in place'community," *AJN The American Journal of Nursing*, vol. 113, no. 1, pp. 68–69, 2013.

[3] M. Skubic, G. Alexander, M. Popescu, M. Rantz, and J. Keller, "A smart home application to eldercare: Current status and lessons learned," *Technology and Health Care*, vol. 17, no. 3, pp. 183–201, 2009.

[4] M. J. Rantz, M. Skubic, S. J. Miller, C. Galambos, G. Alexander, J. Keller, and M. Popescu, "Sensor technology to support aging in place," *Journal of the American Medical Directors Association*, vol. 14, no. 6, pp. 386–391, 2013.

[5] P. Calyam, I. Jahnke, A. Mishra, R. B. Antequera, D. Chemodanov, and M. Skubic, "Toward an eldercare living lab for sensor-based health assessment and physical therapy," *IEEE Cloud Computing*, vol. 4, no. 3, pp. 30–39, 2017.

[6] M. Skubic, H. Jimison, J. Keller, M. Popescu, M. Rantz, J. Kaye, and M. Pavel, "A framework for harmonizing sensor data to support embedded health assessment," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2014, pp. 1747–1751.

[7] A. Jain, M. Popescu, J. Keller, M. Rantz, and B. Markway, "Linguistic summarization of in-home sensor data," in *Journal of biomedical informatics*, vol. 96. IEEE, 2019, p. 103240.

[8] R. Srikant and R. Agrawal, "Mining quantitative association rules in large relational tables," in *Proceedings of the 1996 ACM SIGMOD international conference on Management of data*, 1996, pp. 1–12.

[9] Z. Hai, K. Chang, and Kim, "Implicit feature identification via co-occurrence association rule mining," in *Gelbukh A.F. (eds) Computational Linguistics and Intelligent Text Processing.CICLing 2011. Lecture Notes in Computer Science*, vol. 6608. Springer, Berlin, Heidelberg, 2011, doi:10.1007/978-3-642-19400-9_31.

[10] J. Kacprzyk and S. Zadrozny, "Linguistic summarization of data sets using association rules," in *The 12th IEEE International Conference on Fuzzy Systems, 2003. FUZZ '03., St. Louis, MO, USA*, vol. 1, 2003, pp. 702–707, doi:10.1109/FUZZ.2003.1209449.

[11] C. Galambos, M. Rantz, A. Craver, M. Bongiorno, M. Pelts, A. J. Holik, and J. S. Jun, "Living with intelligent sensors: Older adult and family member perceptions," *CIN: Computers, Informatics, Nursing*, vol. 37, no. 12, pp. 615–627, 2019.

[12] C. Jiao, P. Lyons, A. Zare, L. Rosales, and M. Skubic, "Heart beat characterization from ballistocardiogram signals using extended functions of multiple instances," in *2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2016, pp. 756–760.

[13] M. Enayati, M. Skubic, J. M. Keller, M. Popescu, and N. Z. Farahani, "Sleep posture classification using bed sensor data and neural networks," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 461–465.

[14] B. Y. Su, K. Ho, M. Skubic, and L. Rosales, "Pulse rate estimation using hydraulic bed sensor," in *2012 annual international conference of the IEEE engineering in medicine and biology society*. IEEE, 2012, pp. 2587–2590.

[15] L. Rosales, M. Su, M. Skubic, and K. Ho, "Heart rate monitoring using hydraulic bed sensor ballistocardiogram," in *Journal of Ambient Intelligence and Smart Environments*, 2017, pp. 193–207.

[16] K. Lydon, B. Y. Su, L. Rosales, M. Enayati, K. Ho, M. Rantz, and M. Skubic, "Robust heartbeat detection from in-home ballistocardiogram signals of older adults using a bed sensor," in *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2015, pp. 7175–7179.

[17] T. Banerjee, J. M. Keller, M. Skubic, and E. Stone, "Day or night activity recognition from video using fuzzy clustering techniques," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 3, pp. 483–493, 2014.

[18] R. Wallace, C. Abbott, C. Gibson-Horn, M. Rantz, and M. Skubic, "Metrics from in-home sensor data to assess gait change due to weighted vest therapy," *Smart Health*, vol. 3, pp. 1–19, 2017.

[19] L. Phillips, C. Deroche, M. Rantz, G. Alexander, L. D. M Skubic, C. Abbott, B. Harris, C. Galambos, and R. Koopman, "Using embedded sensors in independent living to predict gait changes and falls," in *JWestern Journal of Nursing Research*, 2016.

[20] P. Newland, A. Salter, A. Flach, L. Flick, F. P. Thomas, E. E. Gulick, M. Rantz, and M. Skubic, "Associations between self-reported symptoms and gait parameters using in-home sensors in persons with multiple sclerosis," *Rehabilitation Nursing Journal*, vol. 45, no. 2, pp. 80–87, 2020.