JNS

CLINICAL ARTICLE

Characterization of patients with idiopathic normal pressure hydrocephalus using natural language processing within an electronic healthcare record system

***Jonathan P. Funnell, MBBS,[1,2] Kawsar Noor,[3,4] Danyal Z. Khan, MRCS,[1,2] Linda D'Antona, MBBS, MD,[2,5] Richard J. B. Dobson, PhD,[3,4,6–8] John G. Hanrahan, MBBS,[1,2] Christopher Hepworth, PhD,[9] Eleanor M. Moncur, MBChB,[2,5] Benjamin M. Thomas, BSc,[10] Lewis Thorne, FRCS(SN),[2] Laurence D. Watkins, FRCS(SN), MD,[2] Simon C. Williams, MBChB,[1,2] Wai Keong Wong, FRCPath, MRCP, PhD,[4,6] Ahmed K. Toma, FRCS(SN), MD,[2,4,5] and Hani J. Marcus, FRCS(SN), PhD[1,2,4,5]**

[1]Wellcome/EPSRC Centre for Interventional and Surgical Sciences, University College London; [2]National Hospital for Neurology and Neurosurgery, London; [3]Institute for Health Informatics, University College London; [4]NIHR Biomedical Research Centre, University College London Hospitals NHS Foundation Trust, London; [5]UCL Queen Square Institute of Neurology, University College London; [6]Health Data Research UK London, University College London; [7]NIHR Biomedical Research Centre, South London and Maudsley NHS Foundation Trust and King's College London; [8]Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London; [9]B. Braun Medical Ltd., Sheffield, United Kingdom; and [10]B. Braun Medical AB, Danderyd, Sweden

OBJECTIVE Idiopathic normal pressure hydrocephalus (iNPH) is an underdiagnosed, progressive, and disabling condition. Early treatment is associated with better outcomes and improved quality of life. In this paper, the authors aimed to identify features associated with patients with iNPH using natural language processing (NLP) to characterize this cohort, with the intention to later target the development of artificial intelligence–driven tools for early detection.

METHODS The electronic health records of patients with shunt-responsive iNPH were retrospectively reviewed using an NLP algorithm. Participants were selected from a prospectively maintained single-center database of patients undergoing CSF diversion for probable iNPH (March 2008–July 2020).

Analysis was conducted on preoperative health records including clinic letters, referrals, and radiology reports accessed through CogStack. Clinical features were extracted from these records as SNOMED CT (Systematized Nomenclature of Medicine Clinical Terms) concepts using a named entity recognition machine learning model.

In the first phase, a base model was generated using unsupervised training on 1 million electronic health records and supervised training with 500 double-annotated documents. The model was fine-tuned to improve accuracy using 300 records from patients with iNPH double annotated by two blinded assessors. Thematic analysis of the concepts identified by the machine learning algorithm was performed, and the frequency and timing of terms were analyzed to describe this patient group.

RESULTS In total, 293 eligible patients responsive to CSF diversion were identified. The median age at CSF diversion was 75 years, with a male predominance (69% male). The algorithm performed with a high degree of precision and recall (F1 score 0.92).

Thematic analysis revealed the most frequently documented symptoms related to mobility, cognitive impairment, and falls or balance. The most frequent comorbidities were related to cardiovascular and hematological problems.

CONCLUSIONS This model demonstrates accurate, automated recognition of iNPH features from medical records. Opportunities for translation include detecting patients with undiagnosed iNPH from primary care records, with the aim to

ABBREVIATIONS AI = artificial intelligence; EHR = electronic health record; iNPH = idiopathic normal pressure hydrocephalus; NLP = natural language processing; SNOMED CT = Systematized Nomenclature of Medicine Clinical Terms.

SUBMITTED May 14, 2022. **ACCEPTED** September 8, 2022.
INCLUDE WHEN CITING Published online November 18, 2022; DOI: 10.3171/2022.9.JNS221095.

* J.P.F. and K.N. contributed equally to this work and share first authorship. A.K.T. and H.J.M. contributed equally to this work and share senior authorship.

ultimately improve outcomes for these patients through artificial intelligence–driven early detection of iNPH and prompt treatment.

https://thejns.org/doi/abs/10.3171/2022.9.JNS221095

KEYWORDS artificial intelligence; cerebrospinal fluid diversion; machine learning; natural language processing; normal pressure hydrocephalus

Idiopathic normal pressure hydrocephalus (iNPH) is a condition of elderly patients that is typically characterized by a triad of gait disturbance, cognitive impairment, and urinary incontinence.[1,2] Without surgical intervention, it is progressive and neurological deterioration may be rapid, resulting in reduced quality of life and life expectancy.[3,4] Prompt treatment is recognized to improve patients' mobility, cognition, urinary problems, quality of life, and dependence on care; however, delays in referrals remain.[5] Additionally, identifying which patients will respond to surgical intervention is difficult without invasive investigations.[6]

A significant barrier to the early detection and referral of patients with iNPH is the overlap between features of iNPH and its mimics. Gait disturbances, cognitive impairment, and urinary incontinence are common among elderly individuals and share features with other neurodegenerative disorders.[7,8] Furthermore, neurodegenerative disorders and iNPH may coexist, adding additional complexity to diagnosis and management.[9,10] Several studies have explored the use of artificial intelligence (AI) in identifying features of iNPH via diagnostic imaging and distinguishing them from mimicking disorders in patients; however, to our knowledge, no group has attempted to detect patients with iNPH using AI-driven analysis of written medical records.[11,12]

Natural language processing (NLP) is a branch of AI that involves training machines to analyze and interpret unstructured written information within its intended context.[13] Within healthcare, NLP is particularly relevant to the interpretation and amalgamation of information from clinical notes, masses of which are stored as unstructured entries within electronic health record (EHR) systems. Extraction of these data using AI confers advantages such as collating large volumes of data for analysis that may otherwise be too labor-intensive to collect, across a range of electronic data sources.

In this study, we aimed to detect and evaluate indicators of iNPH using an NLP machine learning algorithm using clinical information extracted from written clinic notes, outpatient correspondence, and radiology reports. We believe that an application of NLP may be the early detection of iNPH by identifying early features of the disease through automated screening of patients' medical records. The intention of this study is to inform future research and help develop tools to assist clinicians in the detection of probable iNPH for further assessment.

Methods

Compliance With Ethical Standards

This study was registered as part of a service evaluation within University College London Hospitals and approved by the Clinical Governance Committee. Informed consent was not required for this study.

Study Design

This study used NLP to analyze clinical records from patients investigated or treated for iNPH at a tertiary academic neurosurgical center in the United Kingdom (National Hospital for Neurology and Neurosurgery, London). Descriptive analysis of clinical information extracted using an NLP-based machine learning algorithm is presented.

Written EHRs comprising clinical notes, correspondence, and radiology reports were collated for patients undergoing investigation or treatment for iNPH over a consecutive period of 12 years (March 2008–July 2020). Patients were identified from a prospectively maintained clinical database of patients undergoing invasive investigation or treatment for probable iNPH and included if they demonstrated a clinically significant improvement following CSF diversion.

Patients previously investigated or treated at other neurosurgical centers and those with suspected secondary NPH are not included in the database (excluding previous lumbar puncture). Patients for whom EHRs could not be retrieved were excluded.

Demographic information such as age, sex, and ethnicity was extracted from structured data elements within the EHR from the time of CSF diversion. Documents were included up to 5 years prior to the first CSF diversion procedure (lumbar drainage, ventriculoperitoneal shunt, or lumboperitoneal shunt).

Investigation of Suspected iNPH

All patients included in this study were investigated for iNPH according to the routine protocol used by this center. Patients referred to the center with suspected iNPH are evaluated by a consultant neurosurgeon with a subspecialty interest in CSF disorders, and MRI of the brain is performed. Clinical examination including assessment of walking speed and stride length using a 10-m walking test is performed.

Patients with symptoms suggestive of iNPH are offered a trial of extended lumbar CSF drainage over a 72-hour period. Standardized pre– and post–lumbar drainage walking tests are performed and compared, as well as pre– and post–lumbar drainage neuropsychological assessment. An improvement of 10% in walking speed, stride length, or verbal or performance IQ (Wechsler Adult Intelligence Scale) is considered clinically significant and the patient is offered permanent CSF diversion by ventriculoperitoneal shunt.

Model Training and Evaluation

Written clinical information for eligible patients was extracted using an information retrieval platform (CogStack) from the EHR (Epic Systems).[14] This written in-

TABLE 1. Meta-annotation concepts used during the data annotation phase

Meta-Annotation Concept	Option 1	Option 2	Explanation
Negation	No	Yes	Annotated concept is negated/not present, e.g., “[COVID] negative,” where annotation of the written term “[COVID]” suggests infection with COVID-19
Experiencer	Patient	Other	Annotated concept relates to another individual rather than the patient, e.g., “Family history of [ischemic heart disease]” where “[ischemic heart disease]” is annotated
Certainty	Confirmed	Suspected	Annotated concept relates to a diagnosis that is not confirmed, e.g., “MRI scan requested to assess for possible [space-occupying lesion],” where “[space occupying lesion]” is annotated

formation was subsequently inputted into the CogStack Natural Language Processing Platform to make the data available for analysis using NLP tools.[14] In particular, MedCAT, a named entity recognition machine learning model, was used to identify Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) concepts within the notes.[15]

The model was trained in two phases: 1) to construct a base model and 2) to construct a model specifically trained using the medical records of patients with iNPH. In the first phase, the model was trained using unsupervised machine learning pretrained on 1 million EHRs randomly sampled from the hospital-wide EHR. This model was refined by a supervised learning stage in which 500 documents were annotated with SNOMED CT concepts by two independent assessors to create a base model for use across a range of clinical applications.

The second phase was devised with the purpose of training the model using documents derived from the medical records of patients with iNPH. This stage involved supervised learning whereby the base model was used to preannotate 300 documents, which were subsequently validated by two independent blinded reviewers (J.P.F. and D.Z.K.) to improve accuracy using the MedCATTrainer interface.[16] Assessors reviewed existing annotations made by the base model to determine whether the suggested matching SNOMED CT concept matched the intended meaning of the annotated word; they labeled the term as correct or incorrect or chose an alternative concept to best fit the written information. Discrepancies between the two assessors were resolved initially by discussing the conflicting terms, and if agreement could not be reached, a third assessor (H.J.M.) served as the arbitrator. The accuracy of the concept annotations made by the model and interobserver agreement are presented as percentage values.

Meta-annotations were used to refine the accuracy of the written information and assess its relevance to the patient (Table 1). These tags are useful adjuncts when the annotation is an accurate representation of the term’s intended meaning; however, the context limits its relevance to the patient. For example, a diagnosis documented in the notes might relate to a pertinent negative (absence of a finding), a diagnosis given to another individual, or a diagnosis that is not yet confirmed. Terms marked with meta-annotations of “negation = yes,” “experiencer = other,” and “certainty = suspected” were excluded from thematic analysis as they were not considered relevant to the patient.

The joint arbitrated data set generated through the MedCATTrainer interface was thereafter used to train the machine learning model using 50 iterations on a training set, representing 80% of the eligible documents.

A further 20% of eligible documents were reserved as a testing set, which was subsequently assessed using k-fold validation (k = 5). Precision and recall were assessed by calculation of a macro F1 score, where the macro F1 score is the average F1 score for all SNOMED CT concepts identified and the F1 score is the mean of precision and recall.

Thematic Analysis

SNOMED CT concepts identified by the NLP model were extracted, and those labeled as “findings” or “disorder” were included in the analysis. Across training and testing data sets, all extracted concepts were included in thematic analysis.

Concepts occurring fewer than five times over the study period were excluded. Additionally, as mentioned above, concepts with meta-annotations of “negation = yes,” “experiencer = other,” and “certainty = suspected” were excluded because they had limited or no relevance to the patient. Thematic analysis was performed to amalgamate similar concepts, and a grounded theory approach was used to produce a list of features to characterize this cohort of patients.[17]

To identify temporal relationships between themes, concepts were additionally grouped according to the time interval between their notation into the medical record and date of CSF diversion.

The full list of concepts was initially reviewed by the first author (J.P.F.) and provisional themes were identified to group similar concepts. This list was then reviewed and refined by the second author (D.Z.K.), and the final list of themes was agreed on by two authors (J.P.F. and D.Z.K.), with disagreements arbitrated by the senior author (H.J.M.).

An overview of the workflow involved in model training, data set generation, and thematic analysis is presented in Fig. 1.

Results

Demographics

From the iNPH database, 322 consecutive patients were identified, of whom 29 were excluded because their EHRs were not available. Clinical information for the 293 eligible patients was extracted and included in the analy-

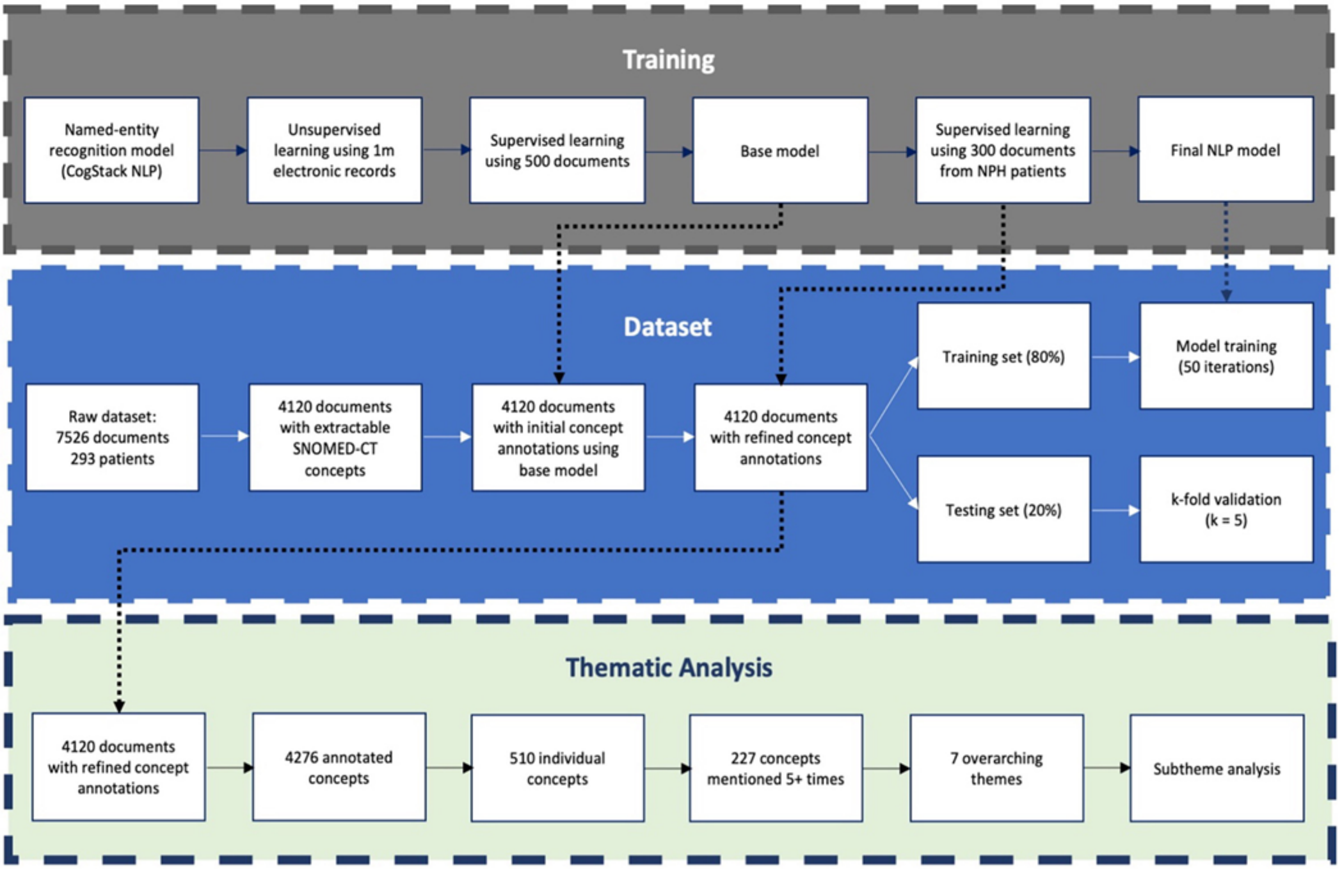

FIG. 1. An overview of the workflow involved in model training, extraction of the data set, and thematic analysis. 1m = 1 million. Figure is available in color online only.

sis. The median age at CSF diversion was 75 years (range 51–93 years), with a male predominance (202 males [69%] and 91 females [31%]). The patients' identified race was predominantly White (205 patients [70%]).

Model Training and Validation

Concept identification using the base model was highly accurate, with the suggested SNOMED CT concept annotated by the model frequently matching the intended meaning. Annotator 1 (J.P.F.) confirmed the accuracy of annotations for 96% of the terms by the base model; annotator 2 (D.Z.K.) confirmed the accuracy for 97% of the terms. Interannotator agreement was equally high, with 97% agreement of terms during the blinded stage. After discussion of discrepancies, agreement on concept accuracy was made on all terms prior to supervised model training.

After refinement with additional supervised annotations, high accuracy of precision and recall was noted with a macro F1 score of 0.92. Because of the infrequency of terms flagged with the meta-annotation "certainty = suspected," no concepts were highlighted by the final model as suspected findings or diagnoses.

Concept Extraction and Thematic Analysis

In total, 7526 documents were extracted, of which 4120 documents (1688 outpatient letters, 921 clinical notes, and 1511 diagnostic imaging reports) included codable SNOMED CT concepts.

Five hundred ten SNOMED CT concepts were extracted from the medical records, of which 227 were mentioned at least five times across the library of records and considered eligible for analysis of the concept theme. This represents analysis of 3863 concept mentions, from a total number of 4276 extracted concept mentions (90.3%).

Seven overarching themes were identified: "comorbidities," "demographics," "patient pathway," "radiological features," "social circumstances," "symptoms and signs," and "treatments." These were each then divided into subthemes. Thirty concepts were excluded from thematic analysis, leaving 197 for inclusion.

Sixty-eight individual concepts were identified as being related to symptoms or clinical signs; these were further subdivided into themes as shown in Table 2. As expected, symptoms and signs relating to recognized signs of iNPH are highly represented among the extracted concepts, with findings related to mobility most prevalent (277 mentions). Among all extracted concepts, the most common individual concept related to falls (59 mentions), followed by reduced mobility (56 mentions).

Concepts representing comorbidities were further di-

TABLE 2. Frequency of concepts extracted for themes relating to symptoms and signs

Symptoms and Signs Theme	Concept Frequency
Mobility	277
Cognitive impairment	138
Falls and balance	107
Bladder or bowel disturbance	93
Pain	88
Level of consciousness	57
Eating and drinking	38
Insomnia	25
Mood	25
Bleeding	21
Dizziness	21
Tremor	18
Weakness	14
Facial flushing	13
Bradykinesia	11
Disability	11
Erythema	11
Tachycardia	10
Nausea	8
Weakness	8
Dysphasia	6
Numbness	5

TABLE 3. Frequency of concepts extracted for themes relating to comorbidities

Comorbidity Theme	Concept Frequency
Cardiovascular	270
Hematological	240
Musculoskeletal	156
Gastrointestinal	111
Neurological	103
Psychiatric	95
Otherwise well	86
Infectious disease	85
Cerebrovascular	84
Renal and urological	71
Systemic or multisystem	62
Hepatopancreatobiliary	55
Respiratory	54
Endocrine	40
Dermatological	28
Ophthalmic	28
Otorhinolaryngological	12
Peripheral vascular	10
Dental	8
Allergy	7

vided by the body system affected (Table 3). The most commonly affected body system of patient comorbidities was cardiovascular (270 concepts), an expected finding in view of the burden of cardiovascular disease among elderly patients in the United Kingdom, especially hypertension (89 mentions). Hematological problems are also significantly represented among the extracted concepts (240 mentions), particularly related to coagulation problems and venous thromboembolism. Among these, commonly cited risks of surgery such as deep vein thrombosis (52 mentions) and pulmonary embolism (50 mentions) were also extracted. The most commonly occurring individual concepts relating to comorbidities included hypertension (89 mentions), stroke (70 mentions), and malignancy (56 mentions).

Concepts relating to patients' social circumstances accounted for 197 mentions, of which 144 related to social support.

Time Series Analysis

As most patients are investigated or treated within a few months of referral, most medical records extracted were collected from the period a few months prior to participants' investigation or intervention for iNPH. This likely reflects the increased correspondence between patients and clinical teams during the preprocedural assessment. Within 1 month prior to the procedure, 65.5% concept mentions were derived, followed by 17.2% between 2 and 3 months prior to the procedure.

When further examining the frequency of the most mentioned symptoms, mentions of all steeply rise in the months leading up to further investigation or intervention. When comparing key features of iNPH (and most common symptom concepts), no theme appeared to emerge sooner than others; however, this may be limited by the dominance of entries in the immediate preprocedural period (Fig. 2).

Discussion

Principal Findings

In this exploratory, proof-of-principle work, we demonstrate AI-driven analysis of medical records of patients with iNPH to characterize this patient cohort. To our knowledge, this is the first study to explore NLP among this group of patients.

Following multiple stages of model training, we demonstrate the automated and accurate detection of clinical features and comorbidities relevant to patients with iNPH. The base model showed high accuracy (96%–97% accuracy of terms) and, following refinement with an additional stage of supervised learning, high precision and recall (F1 score 0.92). This is promising for the further development of NLP-based tools using CogStack in which high accuracy is critical to translation to clinical practice.

In this study, we focused on two groups of clinical variables to characterize patients with iNPH: 1) their symptoms and signs and 2) their comorbidities extracted by CogStack. As expected, the predominant symptoms and signs were related to typical clinical features of iNPH, the most significant of which was mobility. Within "mobility," we subdivided concepts by their relation to it. Interestingly

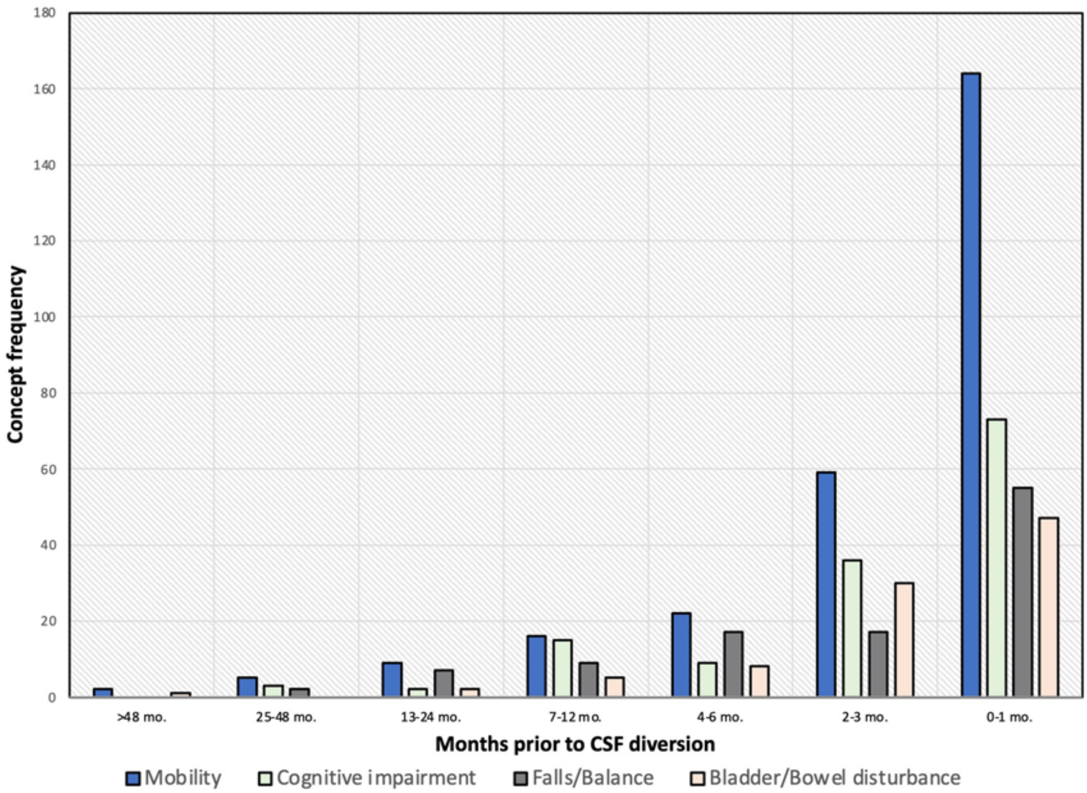

FIG. 2. Changes in concept distribution with time for the four most mentioned themes (relating to signs and symptoms). Figure is available in color online only.

for this group, "gait disturbance" represented the smallest subtheme despite being part of Hakim's triad; instead, clinicians most commonly documented findings related to "reduced mobility." Cognitive impairment was the second most frequent symptoms and signs theme, echoing clinical practice, in which cognitive impairment is less common or recognized than gait disturbance.[18] In addition, concepts relating to "falls and balance" and "bladder or bowel disturbance" were highly represented, as expected for this cohort, albeit less frequently than "mobility" or "cognitive impairment."

The most frequent individual SNOMED CT concept relating to symptoms and signs was "falls," which, while not a cardinal feature of iNPH, are a common consequence of iNPH symptoms. As a result, apart from gait disturbance, balance is considered a separate feature of iNPH in the iNPH symptom scale suggested by Hellström et al.[19]

Comorbidities among patients with iNPH are common. Patients with iNPH tend to be elderly, and with advanced age we more commonly encounter frailty and the complexities of multimorbidity.[20] In this series, we identified cardiovascular comorbidities to be the most significant (by frequency of documentation), driven predominantly by the prevalence of hypertension. This echoes previous experience, particularly among patients older than 80 years, 54% of whom have been previously shown to have hypertension at shunt insertion.[21] This is similar to Finnish cohorts, in which 52% of patients treated for iNPH were noted to have hypertension at the time of shunt insertion.[22]

In contrast to the expected findings of highly prevalent cardiovascular comorbidities, it is perhaps unexpected that hematological problems were so highly represented. This is driven by the high frequency of concepts relating to coagulation and venous thromboembolism, and while these issues are more commonly found among older and immobile individuals, the frequency is higher than we might expect among this group. In contrast, previous case series at our center have described the prevalence of previous pulmonary embolism as 2.56% among iNPH patients, which is substantially less common than the frequency of terms would suggest.[21]

Not all conditions described in medical notes relate to confirmed diagnoses; differential diagnoses and possible risks of procedures are commonly documented. As a result, we expected the model to extract diagnoses that were not confirmed (and therefore not relevant), so a meta-annotation was used to tag these instances during supervised learning. Because of the infrequency of suspected concepts, further training would be necessary to successfully identify diagnoses that are not confirmed, and therefore no suspected terms were detected. Therefore, we suggest that the detected frequency of venous thromboembolism is artificially high, as this is a commonly cited risk associated with surgical interventions when obtaining informed consent. Similarly, while cerebrovascular diseases were commonly extracted, the rate of these diseases found in this study is significantly higher than that from our experience in an earlier case series, in which 6% of patients presented with a history of stroke or transient ischemic attack.[23]

Comparisons to the Literature

Given its predominance in elderly populations, iNPH is predicted to become more prevalent and result in a greater disease burden worldwide, as global population demographics tip toward an aging population. Population-based studies in a Swedish cohort have estimated the prevalence of iNPH among those older than 80 years of age to be as high as 3.7%, yet the incidence of surgical intervention for these patients is low.[24] This suggests that despite the prevalence of iNPH, many patients remain undiagnosed.[25,26] Therefore, sensitive methods for the early detection of iNPH have the potential to confer significant benefits to patients and prevent further deterioration.

There is growing recognition of the utility of big data and AI in the detection of neurological disorders; however, work in iNPH to date has focused on automated detection using AI-driven radiomics.[11,12] Irie et al. demonstrated the effective use of deep learning–enabled analysis of MR images acquired in patients with iNPH or Alzheimer's disease and healthy controls. They demonstrated a high sensitivity and specificity for iNPH (each 91%), which is not significantly different from the sensitivity and specificity found during examination by a radiologist.[11] These techniques are anticipated to be especially clinically useful in the automated analysis of large volumes of MR images, such as those that may be generated by an MRI-based screening program, as has been suggested in a population-based study in Japan.[27] Despite these advances in the analysis of neuroimaging findings, imaging alone is not diagnostic for iNPH and invasive diagnostic tests remain the most effective predictor of shunt response.[6,8,28]

NLP is an exciting opportunity for clinicians since clinical data routinely collected as part of patient care can be rapidly analyzed and interpreted, and much work has been done to translate these data into measures of clinical risk. Within general surgery, Soguero-Ruiz et al. used a bag-of-words model to detect anastomotic leak among patients undergoing surgery for colorectal cancer.[29] With this technique, Soguero-Ruiz et al. demonstrated 100% sensitivity (and 72% specificity) for the detection of postsurgical complications.[29] This highlights the functionality that NLP has in detecting the clinical risks for retrospective audit and research, and raises the question of whether postsurgical complications can be predicted in real time and identified at an earlier stage.

Through the lens of characterizing iNPH, we describe the use of CogStack, an information retrieval and NLP platform, which has been used to calculate real-time clinical risks. Oliver et al. demonstrated the use of CogStack to analyze the EHRs of patients accessing psychiatric care in South London and flag individuals at high risk of psychosis to clinicians for review.[30] In this feasibility study, Oliver et al. collected structured clinical data points and used these as part of a risk calculator (therefore not using NLP), demonstrating that integrating flags of high risk within the EHR is feasible.

The aim of our group was to unify these two concepts and apply them to iNPH. Principally, these concepts are 1) NLP-enabled identification of clinical concepts relevant to iNPH and 2) the automated detection and flagging of at-risk patients for review.

Strengths and Limitations

In this exploratory work, we describe the use of NLP-based case note analysis for patients with iNPH, the first of its kind, to our knowledge. We intended to demonstrate the effective use of automated clinical concept extraction among iNPH patients achieved through the development of a model with high precision and recall. One main strength of our model is that it derives from multiple stages of training with a large data set, iteratively refining the model's ability to interpret clinical features among patients with iNPH.

Although our data set is large, the nature of the records (retrospective, single center) makes the model vulnerable to overfitting, impacting features extracted and themes described. Concepts described in our thematic analysis are captured from patients who share multiple common variables (diagnosis, clinicians, referral pathway, and treatment algorithm). As a result, the pattern of writing is likely to be similar across patients, not only because they share clinical features but also because the records are derived from the same authors. For this reason, future research should analyze primary care records for patients with iNPH that are likely to be more heterogeneous and thus may be more useful in identifying themes not previously conventionally associated with iNPH.

A further limitation of this data set is that few entries were written more than 3 months prior to the procedure, as most patients do not wait a prolonged length of time from referral to CSF diversion. This guides our focus to primary care in which automated screening tools are likely to be more useful in detecting undiagnosed cases of iNPH.

Conclusions

Our findings demonstrate the development of a model to extract features of iNPH from EHRs with high accuracy and describe clinical features of iNPH through the lens of NLP.

Although this work cannot yet distinguish between patients with and without shunt-responsive iNPH, this proof-of-principle study demonstrates the extraction of relevant concepts for these patients and provides the basis for further research. In the future, we aim to evaluate the algorithm in its ability to distinguish patients with iNPH and those with mimicking conditions at an earlier stage in primary care. We also aim to evaluate the effectiveness of NLP in identifying these patients compared with that of AI-enabled radiomics.

The overall aim of our future work is to predict outcomes following CSF diversion and work toward the automated early detection of iNPH among undiagnosed individuals, flagging patients to streamline access to lumbar drainage and further assessment.

Acknowledgments

This study was supported by the National Institute for Health Research (NIHR) University College London (UCL) Hospitals

(UCLH) Biomedical Research Centre (BRC), in particular by the NIHR UCLH/UCL BRC Clinical and Research Informatics Unit.

References

1. Williams MA, Relkin NR. Diagnosis and management of idiopathic normal-pressure hydrocephalus. *Neurol Clin Pract*. 2013;3(5):375-385.
2. Mori E, Ishikawa M, Kato T, et al. Guidelines for management of idiopathic normal pressure hydrocephalus: second edition. *Neurol Med Chir (Tokyo)*. 2012;52(11):775-809.
3. Toma AK, Stapleton S, Papadopoulos MC, Kitchen ND, Watkins LD. Natural history of idiopathic normal-pressure hydrocephalus. *Neurosurg Rev*. 2011;34(4):433-439.
4. Petersen J, Hellström P, Wikkelsø C, Lundgren-Nilsson A. Improvement in social function and health-related quality of life after shunt surgery for idiopathic normal-pressure hydrocephalus. *J Neurosurg*. 2014;121(4):776-784.
5. Toma AK, Papadopoulos MC, Stapleton S, Kitchen ND, Watkins LD. Systematic review of the outcome of shunt surgery in idiopathic normal-pressure hydrocephalus. *Acta Neurochir (Wien)*. 2013;155(10):1977-1980.
6. Thavarajasingam SG, El-Khatib M, Rea M, et al. Clinical predictors of shunt response in the diagnosis and treatment of idiopathic normal pressure hydrocephalus: a systematic review and meta-analysis. *Acta Neurochir (Wien)*. 2021;163(10):2641-2672.
7. Jeppsson A, Wikkelsö C, Blennow K, et al. CSF biomarkers distinguish idiopathic normal pressure hydrocephalus from its mimics. *J Neurol Neurosurg Psychiatry*. 2019;90(10):1117-1123.
8. Williams MA, Malm J. Diagnosis and treatment of idiopathic normal pressure hydrocephalus. *Continuum (Minneap Minn)*. 2016;22(2 Dementia):579-599.
9. Pomeraniec IJ, Bond AE, Lopes MB, Jane JAS Sr. Concurrent Alzheimer's pathology in patients with clinical normal pressure hydrocephalus: correlation of high-volume lumbar puncture results, cortical brain biopsies, and outcomes. *J Neurosurg*. 2016;124(2):382-388.
10. Müller-Schmitz K, Krasavina-Loka N, Yardimci T, et al. Normal pressure hydrocephalus associated with Alzheimer's disease. *Ann Neurol*. 2020;88(4):703-711.
11. Irie R, Otsuka Y, Hagiwara A, et al. A novel deep learning approach with a 3D convolutional ladder network for differential diagnosis of idiopathic normal pressure hydrocephalus and Alzheimer's disease. *Magn Reson Med Sci*. 2020;19(4):351-358.
12. Sotoudeh H, Sadaatpour Z, Rezaei A, et al. The role of machine learning and radiomics for treatment response prediction in idiopathic normal pressure hydrocephalus. *Cureus*. 2021;13(10):e18497.
13. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg*. 2018;268(1):70-76.
14. Noor K, Roguski L, Bai X, et al. Deployment of a free-text analytics platform at a UK National Health Service research hospital: CogStack at University College London Hospitals. *JMIR Med Inform*. 2022;10(8):e38122.
15. Kraljevic Z, Searle T, Shek A, et al. Multi-domain clinical natural language processing with MedCAT: the Medical Concept Annotation Toolkit. *Artif Intell Med*. 2021;117:102083.
16. Searle T, Kraljevic Z, Bendayan R, Bean D, Dobson R. MedCATTrainer: a biomedical free text annotation interface with active learning and research use case specific customisation. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations*. Association for Computational Linguistics; 2019;139-144.
17. Chapman AL, Hadfield M, Chapman CJ. Qualitative research in healthcare: an introduction to grounded theory using thematic analysis. *J R Coll Physicians Edinb*. 2015;45(3):201-205.
18. Relkin N, Marmarou A, Klinge P, Bergsneider M, Black PM. Diagnosing idiopathic normal-pressure hydrocephalus. *Neurosurgery*. 2005;57(3 suppl):S4-S16, ii-v.
19. Hellström P, Klinge P, Tans J, Wikkelsø C. A new scale for assessment of severity and outcome in iNPH. *Acta Neurol Scand*. 2012;126(4):229-237.
20. Soon SXY, Kumar AA, Tan AJL, et al. The impact of multimorbidity burden, frailty risk scoring, and 3-directional morphological indices vs. testing for CSF responsiveness in normal pressure hydrocephalus. *Front Neurosci*. 2021;15:751145.
21. Thompson SD, Shand Smith JD, Khan AA, Luoma AMV, Toma AK, Watkins LD. Shunting of the over 80s in normal pressure hydrocephalus. *Acta Neurochir (Wien)*. 2017;159(6):987-994.
22. Pyykkö OT, Nerg O, Niskasaari HM, et al. Incidence, comorbidities, and mortality in idiopathic normal pressure hydrocephalus. *World Neurosurg*. 2018;112:e624-e631.
23. D'Antona L, Blamey SC, Craven CL, et al. Early postoperative outcomes of normal pressure hydrocephalus: results of a service evaluation. *J Neurosurg Anesthesiol*. 2021;33(3):247-253.
24. Andersson J, Rosell M, Kockum K, Lilja-Lund O, Söderström L, Laurell K. Prevalence of idiopathic normal pressure hydrocephalus: a prospective, population-based study. *PLoS One*. 2019;14(5):e0217705.
25. Sundström N, Malm J, Laurell K, et al. Incidence and outcome of surgery for adult hydrocephalus patients in Sweden. *Br J Neurosurg*. 2017;31(1):21-27.
26. Brean A, Fredø HL, Sollid S, Müller T, Sundstrøm T, Eide PK. Five-year incidence of surgery for idiopathic normal pressure hydrocephalus in Norway. *Acta Neurol Scand*. 2009;120(5):314-316.
27. Iseki C, Takahashi Y, Wada M, Kawanami T, Adachi M, Kato T. Incidence of idiopathic normal pressure hydrocephalus (iNPH): a 10-year follow-up study of a rural community in Japan. *J Neurol Sci*. 2014;339(1-2):108-112.
28. Craven CL, Toma AK, Mostafa T, Patel N, Watkins LD. The predictive value of DESH for shunt responsiveness in idiopathic normal pressure hydrocephalus. *J Clin Neurosci*. 2016;34:294-298.
29. Soguero-Ruiz C, Hindberg K, Rojo-Alvarez JL, et al. Support vector feature selection for early detection of anastomosis leakage from bag-of-words in electronic health records. *IEEE J Biomed Health Inform*. 2016;20(5):1404-1415.
30. Oliver D, Spada G, Colling C, et al. Real-world implementation of precision psychiatry: transdiagnostic risk calculator for the automatic detection of individuals at-risk of psychosis. *Schizophr Res*. 2021;227:52-60.

Disclosures

L.D.W. reports receiving honoraria from B. Braun and Medtronic; and being on the speakers bureau for Codman. J.P.F., D.Z.K., J.G.H., S.C.W., and H.J.M. are supported by the Wellcome (203145Z/16/Z)/EPSRC (NS/A000050/1) Centre for Interventional and Surgical Sciences, University College London. D.Z.K. and L.D. are supported by NIHR Academic Clinical Fellowships. K.N., W.K.W., A.K.T., and H.J.M. are supported by the NIHR Biomedical Research Centre, University College London. R.J.B.D. is supported by the NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London; Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council,

Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation, and Wellcome Trust; UK Research and Innovation London Medical Imaging & Artificial Intelligence Centre for Value Based Healthcare; and NIHR Applied Research Collaboration South London (NIHR ARC South London) at King's College Hospital NHS Foundation Trust.

Author Contributions

Conception and design: Funnell, Noor, Khan, Dobson, Hepworth, Thomas, Thorne, Watkins, Wong, Toma, Marcus. Acquisition of data: Funnell, Noor, Khan, D'Antona, Moncur, Marcus. Analysis and interpretation of data: Funnell, Noor, Khan, Moncur, Toma, Marcus. Drafting the article: Funnell, Noor, Khan, Hanrahan, Williams, Toma, Marcus. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Funnell. Statistical analysis: Funnell, Noor, Khan. Administrative/technical/material support: Noor, Dobson, Wong. Study supervision: Dobson, Thorne, Watkins, Wong, Toma, Marcus.

Supplemental Information

Previous Presentations

Portions of this work were presented in abstract and oral presentation form at the 2022 Spring Meeting of the Society of British Neurological Surgeons, Cardiff, UK, March 31, 2022.

Correspondence

Jonathan P. Funnell: University College London, United Kingdom. jonathan.funnell.13@ucl.ac.uk.