Received: 2 May 2023 | Revised: 30 June 2023 | Accepted: 7 July 2023

DOI: 10.1002/dad2.12464

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

SHORT REPORT

Dimensional clinical phenotyping using *post-mortem* brain donor medical records: *post-mortem* RDoC profiling is associated with Alzheimer's disease neuropathology

Jonathan Vogelgsang[1] | **Shu Dan**[1] | **Anna P. Lally**[1] | **Michael Chatigny**[1,2] | **Sangeetha Vempati**[1] | **Joshua Abston**[1] | **Peter T. Durning**[1] | **Derek H. Oakley**[2,3] | **Thomas H. McCoy**[4] | **Torsten Klengel**[1,2] | **Sabina Berretta**[1,2]

[1]Department of Psychiatry, McLean Hospital, Harvard Medical School, Belmont, Massachusetts, USA

[2]Harvard Brain Tissue Resource Center, McLean Hospital, Harvard Medical School, Belmont, Massachusetts, USA

[3]Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA

[4]Department of Psychiatry and Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA

Correspondence
Jonathan Vogelgsang and Sabina Berretta, Department of Psychiatry, McLean Hospital, Harvard Medical School, 115 Mill Street, Belmont, MA 02478, USA.
E-mail: jvogelgsang@mclean.harvard.edu; sberretta@mclean.harvard.edu

Funding information
National Institute of Mental Health, Grant/Award Numbers: R01MH120991, R01MH116270, P50NM119467; German Research Foundation (DFG), Grant/Award Number: 413501650; ERA-NET Neuron, Grant/Award Number: 01EW2003; NICHD, Grant/Award Number: R01HD102974; NIA, Grant/Award Numbers: R01AG070704, 5P50MH115874, 5R01MH120991

Abstract

Introduction: Transdiagnostic dimensional phenotypes are essential to investigate the relationship between continuous symptom dimensions and pathological changes. This is a fundamental challenge to *post-mortem* work, as assessments of phenotypic concepts need to rely on existing records.

Methods: We adapted well-validated methodologies to compute National Institute of Mental Health Research Domain Criteria (RDoC) scores using natural language processing (NLP) from electronic health records (EHRs) obtained from *post-mortem* brain donors and tested whether cognitive domain scores were associated with Alzheimer's disease neuropathological measures.

Results: Our results confirm an association of EHR-derived cognitive scores with neuropathological findings. Notably, higher neuropathological load, particularly neuritic plaques, was associated with higher cognitive burden scores in the frontal (ß = 0.38, $P = 0.0004$), parietal (ß = 0.35, $P = 0.0008$), temporal (ß = 0.37, $P = 0.0004$) and occipital (ß = 0.37, $P = 0.0003$) lobes.

Discussion: This proof-of-concept study supports the validity of NLP-based methodologies to obtain quantitative measures of RDoC clinical domains from *post-mortem* EHR. The associations may accelerate *post-mortem* brain research beyond classical case–control designs.

KEYWORDS
biological specimen banks, dementia, medical informatics, natural language processing, neuropathology

1 | INTRODUCTION

Human *post-mortem* brain research provides a critical link between in vitro studies, animal models, and human clinical studies on brain disorders. Its unique contribution is access to the human brain at the cellular and molecular levels. However, *post-mortem* brain research also has some limitations, including the extent of available information regarding donor demographics and, particularly, clinical phenotypes.[1–5] This latter is particularly challenging. In most cases, categorical diagnoses are derived using information gleaned from

Jonathan S. Vogelgsang, Shu Dan, Torsten Klengel, and Sabina Berretta denotes equal contribution.

medical records, the legal next of kin, and neuropathological assessments. This approach is widely used[6,7] and is the backbone of innumerable ground-breaking findings on the pathophysiology of a broad range of brain disorders.[8] In addition to its inherent categorical framework, other limitations, shared by the majority of human studies in this field, are related to clinical phenotypic heterogeneity and overlapping comorbidities typical of many brain disorders and challenges to account for the course of disease over time. These limitations point to the need to bring human brain *post-mortem* investigations in line with current transdiagnostic frameworks[9–14] so that the underlying pathophysiology of brain disorders can be interpreted in a more nuanced manner in the context of transdiagnostic dimensional phenotypes.[15–18] Dimensional clinical phenotypes find support in compelling evidence for substantial overlap of genetic risk as well as clinical, pathophysiological, and pharmacological features across categorical brain disorders.[9–14,19,20]

Recent progress in computational sciences allows for the analysis of health-related data through machine learning, including the analysis of electronic health records (EHR) for comprehensive clinical phenotyping beyond a given diagnosis. Previous literature has focused on using machine learning methods, such as autoencoders and convolutional and recurrent neural networks, to read clinically relevant texts and predict clinical outcomes including readmission rate, risk classifications, discharge timeline, treatment outcome, as well as subphenotyping.[21–23] Among these deep learning approaches, natural language processing (NLP)—software designed to extract information from human-authored narrative-free text—has been widely applied to the medical field to profile various brain disorders and symptoms through algorithms ranging from speech recognition to syntax and sentiment analysis.[24] In the clinical setting, this methodology has been validated against expert annotation, formal cognitive testing, and clinical prediction tasks.[25–31] To our knowledge, it has not yet been applied to the health records from *post-mortem* brain donors and used in combination with neuropathological readouts.

Potential frameworks that can be applied to EHR for multidimensional clinical phenotyping include the Research Domain Criteria (RDoC) and the Hierarchical Taxonomy of Psychopathology (HiTOP).[9,10] The National Institute of Mental Health (NIMH) developed the RDoC framework, a clinical domain-based approach with each domain designed to capture a spectrum of symptoms rooted in brain circuits and biology.[11–14] The domain-specific symptom burden can be estimated from patient medical records using NLP.[28,32,33] RDoC symptom burdens estimated from medical records by NLP have been associated with genetic variants and clinical outcomes including suicide, hospital use, new dementia diagnosis, and progression from dementia diagnosis to death.[34–36]

We put forward that application of NLP-based methodologies to human brain *post-mortem* studies may represent a significant step toward a more current and translatable interpretation of molecular and cellular read-outs in the context of transdiagnostic clinical domains and symptom constructs. As a first step toward assessing the feasibility and validity of this approach, we focused on Alzheimer's disease (AD), a disease with distinct symptoms and well-established neuropathological hallmarks. These arise from two dominant protein pathologies: amyloid beta (Aβ), forming extracellular Aβ aggregates known as amyloid or senile plaques and tau, forming intracellular tau accumulations known as neurofibrillary tangles (NFT) and dystrophic neurites (DNs). Neuritic plaques (NP), formed by Aβ plaques containing DNs, are considered a pathologic hallmark of AD.[37] Clinically, AD is characterized by impaired cognition including deficits of amnestic and non-amnestic memory, judgment, and reasoning, as well as impaired visuospatial and language functions.[38–40]

The aim of this study is to provide conceptual evidence for the use of NLP on *post-mortem* brain donor health records and the association of NLP-derived dimensional RDoC phenotypes with AD hallmark neuropathology. In line with prior evidence for an association of cognitive decline and NP burden and as a proof-of-concept study, we hypothesize that NLP-derived scores for the RDoC cognitive domain are associated with NP load.

RESEARCH IN CONTEXT

1. Systematic review: The authors reviewed the most recent literature on PubMed to review the current state of natural language processing–based clinical phenotyping of electronic health records. Although several algorithms have been published in the past, none has been applied to medical records from *post-mortem* brain bank collections.
2. Interpretation: This study proposes the use of natural language-based methods to better characterize *post-mortem* brain donors. Our findings are consistent with published clinicopathological correlation studies combining psychometric in vivo testing with *post-mortem* neuropathological evaluation.
3. Future directions: The used algorithm should be further tested; particularly, the other domains should be validated using clinical cohorts with in-depth phenotypic characterization.

2 | MATERIAL AND METHODS

2.1 | Study cohort

We selected 92 donors (46 males, 46 females, mean age 81.9 years, standard deviation 9.47, range: 57–98 years) from the Harvard Brain and Tissue Resource Center, National Institutes of Health (NIH) NeuroBioBank (HBTRC/NBB) with Braak stages between 0 and 6. *Post-mortem* clinical and neuropathological evaluation was performed by an experienced team of a neuropathologist and two psychiatrists. Briefly, health records were independently reviewed by two clinicians while a neuropathologist created a neuropathological report.

Biweekly consensus meetings are held with all clinical reviewers, neuropathologists, and other team members from various backgrounds, including clinical psychology, medicine, molecular biology, and *post-mortem* brain research. During these meetings, the clinical cases are presented with the neuropathologist reviewing the macroscopic, histological, and molecular findings. Disagreements or open questions are discussed by the entire team to assign a consensus diagnosis. For all cases included in this study, clinical and neuropathological diagnoses were in line, and all neuropathological reports were created by the same neuropathologist. The neuropathological report included gross examination, macroscopic and microscopic assessment of a standard set of brain regions, as well as a semiquantitative neuritic plaque rating. Several features, including neuronal loss and presence of NFTs and NPs were assessed in the frontal (Brodmann area [BA] 3/2/1, 4, 9, 46), parietal (BA 39, 40), temporal (hippocampal formation with lateral geniculate body and tail of caudate nucleus, entorhinal cortex, and anterior hippocampus), and occipital (BA 17, 18/19) cortex. Besides AD-typical neuropathological observations, donors showed age-related vascular changes and incidental/secondary concomitant neuropathological findings, including five cases with Lewy bodies, six cases with hippocampal sclerosis, and six cases with argyrophilic grain disease. Medical records were carefully reviewed. No disease-specific symptoms (e.g., hallucinations related to Lewy bodies) were identified. Thus, these findings were considered insignificant for the analysis. Clinically, donors were either neurotypical or diagnosed with AD/dementia without additional psychiatric or neurological diseases present.

In addition, the HBTRC collects extensive demographic data and clinical data included in health records and a questionnaire completed by the legal next of kin.[41] All medical records available for each case were digitized.

2.2 | Ethics statement

Tissue and medical records have been collected under the HBTRC institutional review board protocol 2015P002028 (McLean/Mass General Brigham Institutional Review Board). Data about brain donors are made available to investigators in de-identified form according to Health Insurance Portability and Accountability Act regulations.

2.3 | Consent statement

Formal informed consent to donate the brain and related samples for research is obtained after death from the legal next of kin and documented in writing.

2.4 | Processing of digital records and scoring clinical text for cognition symptom burden

All available medical notes and records for each donor, regardless of time of creation or content, were scanned and transformed into a text file. Health records included documents covering the psychiatric or neurological status of the donor, documents pertaining to internal medicine, surgery, or other treatments, as well as administrative information. The present study used a previously described and validated NLP algorithm for quantifying estimated cognitive symptoms from narrative clinical text.[28] In brief, this method relies on recognizing a pre-specified set of symptom-related terms within the available records. The term list was developed through an iterative process of refinement seeded with lists of terms developed by a group of clinical experts, including the NIMH Research Domain Criteria Working Group. That seed was subsequently expanded through unsupervised machine learning to enhance coverage of the clinical lexicon.[28] The final cognitive symptom score is the proportion of terms that appear in any given note. The tool is implemented as freely available code for online download, including the full list of tokens as described in the initial validation publication (https://github.com/thmccoy/CQH-Dimensional-Phenotyper).[28] Briefly, the code searches for tokens included in the medical records to calculate domain-specific symptom burdens. Importantly, this tool was not trained to predict nor fitted against any particular outcome; rather, it was developed without regard to any particular categorical diagnosis or outcome, aiming instead to directly capture dimensions of neuropsychiatric symptomatology as dimensional rather than categorical.[11] Similarly, and as in prior work applying this NLP approach to dementia and general medical records across diagnosis, the algorithm was not trained, fitted, calibrated, modified, or otherwise biased toward the samples reported in this paper.[35,42]

2.5 | Analysis and data availability

The demographic characteristics of the cohort were summarized using univariate summary statistics. The primary analysis assessed the strength of association between cognitive symptoms and neuropathological findings of plaques by lobe. For each scanning method, we regressed the NLP-derived cognitive symptom burden score on the NP load of each lobe of the brain, controlling for age and sex, in a random intercept model. All analysis used R v4.2.2. Raw measurements are stored at the primary study site and can be provided upon request.

3 | RESULTS

Mixed effects models, adjusted for age and sex, combining NLP data with neuropathology results showed that higher NP load is associated with higher cognitive burden scores in the frontal (ß = 0.38, $P = 0.0004$), parietal (ß = 0.35, $P = 0.0008$), temporal (ß = 0.37, $P = 0.0004$), and occipital (ß = 0.37, $P = 0.0003$) lobes (Table 1). A second confirmatory regression model comparing cognitive burden and Braak stage, controlling for age and sex, also showed a significant association between RDoC cognition and B&B stage of ß = 0.35 (95% confidence interval: 0.16–0.54, *P*-value = 0.0005).

TABLE 1 Association between neuritic plaque load and cognitive symptom burden in mixed effect model regressing cognitive burden score on neuritic plaque load controlling for age and sex.

	ß	ß 95% CI	*P*-value
Frontal NP	0.38	0.17–0.58	0.0004
Parietal NP	0.35	0.15–0.55	0.0008
Temporal NP	0.37	0.17–0.56	0.0004
Occipital NP	0.37	0.17–0.57	0.0003

Abbreviations: CI, confidence interval; NP, neuritic plaque.

4 | DISCUSSION

NLP methodologies, designed to obtain multidimensional phenotypic fingerprints from EHRs based on the NIMH RDoC framework, have been well validated through expert testing and annotation.[25–31] Our goal was to validate the use of this approach in the context of human *post-mortem* studies. Our study replicates well-established associations of cognitive decline and AD pathology, showing a strong significant correlation between pathological hallmarks of AD, that is, NP density and Braak stages, and cognitive deficits according to the RDoC cognitive domain. Thus, these results offer a first proof of concept in support of the application of RDoC-based NLP algorithms in human *post-mortem* studies of AD and the potential of these methods for more detailed analyses beyond categorical case–control designs in *post-mortem* research in general.

A potential limitation of these studies is that the donor cohort was not followed longitudinally during the course of the disease so that we were not able to assess our results against clinical scales—health records and information from the donors' families were obtained *post-mortem*. Thus, clinical information on donors may vary in quality and quantity and might be biased toward donors with a higher burden of neuropsychiatric symptoms. This is a common challenge in *post-mortem* studies, limiting our options for regressing NLP data against an accurate interpretation of clinical assessments. To address this limitation, AD-related pathological load was used in these studies as a stand-in for cognitive impairment. Although the precise relationship between aspects of AD pathology and cognitive impairment may still be under investigation, compelling evidence shows that the severity of NP and NFT load is a strong predictor of cognitive decline.[43–45] Conversely, the validity of cognitive measures derived by the NLP algorithms used for these studies is supported by extensive studies in much larger cohorts of live subjects.[28,31,35] This methodology has proven to have robust transdiagnostic predictive validity against genetic correlates as well as a broad range of clinical constructs, including agitation, risk of dementia, and risk of suicide,[29,36] supporting its capacity to gain traction on difficult neuropsychiatric problems.

Transdiagnostic dimensional phenotypes are essential to investigations of the relationships between continuous symptom dimensions and cellular/molecular changes in brain disorders. In response to this challenge, efforts over the last decade have focused on overcoming the limitations inherent to categorical diagnostic approaches by establishing dimensional models based on neurobiological or behavioral phenotypes such as RDoC or HiTOP. Although these efforts are not without controversy,[46] dimensional phenotyping across diagnostic entities is a critical tool needed to understand the cell-level patterns of molecular changes underlying clinical domains and symptom constructs in brain disorders. Furthermore, these approaches are needed to investigate urgent questions on the relationships between specific symptoms and underlying pathophysiological mechanisms across disorders. For instance, cognitive impairment, anxiety, and depression are shared by many brain disorders, from dementias to major depression and schizophrenia, and each is largely treated using the same pharmacological approaches. The underlying, often implicit, assumption that similar molecular, cellular, and neural circuit pathology underlies each of these symptoms across these disorders may be plausible, but as yet poorly tested. Arguably, these considerations apply to the neuropathology of dementias, as hallmark pathology, such as the impact of proteinopathies affecting tau, Aβ, and α-synuclein, has been tested against cognition while largely neglecting co-occurring symptoms. Categorical diagnostic approaches, encompassing heterogeneous symptoms under one diagnosis, are not sufficiently nuanced to address these questions.

As the NIMH RDoC framework was not specifically developed for AD or other dementias but for neuropsychiatric disorders in general, RDoC cognition might not reflect identical symptom constructs and should be interpreted with caution. However, McCoy et al. showed in large clinical cohorts that RDoC cognition, as measured with the same code used for this study, is associated with a clinical conversion to dementia.[35] Thus, we believe that this approach contains relevant information related to clinically significant cognitive decline in AD and other dementias.

In conclusion, this proof-of-concept study supports the validity of NLP approaches to extract dimensional clinical phenotypic data from health records obtained *post-mortem* from brain donors. Analyses were limited to the cognitive dimension so that hallmark AD pathology could be used as a well-established predictor of cognitive impairment. Ongoing efforts are focused on the application of these approaches to a broader range of clinical domains and brain disorders.[35]

ACKNOWLEDGMENTS

Brain tissue samples included in these studies were obtained from the Harvard Brain and Tissue Resource Center, an NIH NeuroBioBank site. The authors express gratitude to all brain donors and their families for their generosity. This study was funded by the National Institute of Mental Health (R01MH120991, R01MH116270). The sponsors had no role in the study design, writing of the report, or data collection, analysis, or interpretation. All authors had full access to the data and decided to submit for publication jointly. JV received research funding from the German Research Foundation (DFG), Eric Dorris Memorial Fund, and P50 NM119467. THM received research funding from the Stanley Center at the Broad Institute, the Brain and Behavior Research Foundation, National Institute of Mental Health, National Human Genome Research Institute Home, and Telefonica Alfa. TK received support from ERA-NET Neuron 01EW2003, NICHD R01HD102974, and NIA

R01AG070704. SB received support from NIMH 5P50MH115874 and NIMH 5R01MH120991.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Jonathan Vogelgsang https://orcid.org/0000-0001-9326-8193

REFERENCES

1. McCullumsmith RE, Hammond JH, Shan D, Meador-Woodruff JH. Postmortem brain: an underutilized substrate for studying severe mental illness. *Neuropsychopharmacol*. 2014;39:65-87. doi:10.1038/npp.2013.239
2. Lewis DA. The human brain revisited: opportunities and challenges in postmortem studies of psychiatric disorders. *Neuropsychopharmacol*. 2002;26:143-154. doi:10.1016/s0893-133x(01)00393-1
3. Gomez-Nicola D, Boche D. Post-mortem analysis of neuroinflammatory changes in human Alzheimer's disease. *Alzheimer's Res Ther*. 2015;7:42. doi:10.1186/s13195-015-0126-1
4. Sullivan K, Pantazopoulos H, Liebson E, et al. Chapter 14 What can we learn about brain donors? *Handb Clin Neurology*. 2018;150:181-196. doi:10.1016/b978-0-444-63639-3.00014-1
5. Berretta S, Heckers S, Benes FM. Searching human brain for mechanisms of psychiatric disorders. Implications for studies on schizophrenia. *Schizophr Res*. 2015;167:91-97. doi:10.1016/j.schres.2014.10.019
6. Levey DF, Stein MB, Wendt FR, et al. Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nat Neurosci*. 2021;24:954-963. doi:10.1038/s41593-021-00860-2
7. Gelernter J, Sun N, Polimanti R, et al. Genome-wide association study of post-traumatic stress disorder reexperiencing symptoms in >165,000 US veterans. *Nat Neurosci*. 2019;22:1394-1401. doi:10.1038/s41593-019-0447-7
8. Wang D, Liu S, Warrell J, et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science*. 2018;362:1266. doi:10.1126/science.aat8464
9. Dalgleish T, Black M, Johnston D, Bevan A. Transdiagnostic approaches to mental health problems: current status and future directions. *J Consult Clin Psych*. 2020;88:179-195. doi:10.1037/ccp0000482
10. Hengartner MP, Lehmann SN. Why psychiatric research must abandon traditional diagnostic classification and adopt a fully dimensional scope: two solutions to a persistent problem. *Frontiers Psychiatry*. 2017;8:101. doi:10.3389/fpsyt.2017.00101
11. Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiat*. 2010;167:748-751. doi:10.1176/appi.ajp.2010.09091379
12. Cuthbert BN. The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry*. 2014;13:28-35. doi:10.1002/wps.20087
13. Insel TR, Cuthbert BN. Medicine. Brain disorders? Precisely. *Science*. 2015;348:499-500. doi:10.1126/science.aab2358
14. Carcone D, Ruocco AC. Six years of research on the national institute of mental health's research domain criteria (RDoC) initiative: a systematic review. *Front Cell Neurosci*. 2017;11:46. doi:10.3389/fncel.2017.00046
15. Reininghaus U, Böhnke JR, Chavez-Baldini U, et al. Transdiagnostic dimensions of psychosis in the bipolar-schizophrenia network on intermediate phenotypes (B-SNIP). *World Psychiatry*. 2019;18:67-76. doi:10.1002/wps.20607
16. Parkes L, Tiego J, Aquino K, et al. Transdiagnostic variations in impulsivity and compulsivity in obsessive-compulsive disorder and gambling disorder correlate with effective connectivity in cortical-striatal-thalamic-cortical circuits. *Neuroimage*. 2019;202:116070. doi:10.1016/j.neuroimage.2019.116070
17. Grisanzio KA, Goldstein-Piekarski AN, Wang MY, Ahmed APR, Samara Z, Williams LM. Transdiagnostic symptom clusters and associations with brain, behavior, and daily function in mood, anxiety, and trauma disorders. *Jama Psychiat*. 2017;75:201. doi:10.1001/jamapsychiatry.2017.3951
18. Cowan HR, Mittal VA. Transdiagnostic dimensions of psychiatric comorbidity in individuals at clinical high risk for psychosis: a preliminary study informed by HiTOP. *Frontiers Psychiatry*. 2021;11:614710. doi:10.3389/fpsyt.2020.614710
19. Romero C, Werme J, Jansen PR, et al. Exploring the genetic overlap between twelve psychiatric disorders. *Nat Genet*. 2022;54:1795-1802. doi:10.1038/s41588-022-01245-2
20. Hindley G, Frei O, Shadrin AA, et al. Charting the landscape of genetic overlap between mental disorders and related traits beyond genetic correlation. *Am J Psychiat*. 2022;179:833-843. doi:10.1176/appi.ajp.21101051
21. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. *Npj Digit Med*. 2018;1:18. doi:10.1038/s41746-018-0029-1
22. Solares JRA, Raimondi FED, Zhu Y, et al. Deep learning for electronic health records: a comparative review of multiple deep neural architectures. *J Biomed Inform*. 2020;101:103337. doi:10.1016/j.jbi.2019.103337
23. Lage I, Jr THM, Perlis RH, Doshi-Velez F. Efficiently identifying individuals at high risk for treatment resistance in major depressive disorder using electronic health records. *J Affect Disord*. 2022;306:254-259. doi:10.1016/j.jad.2022.02.046
24. Hecker P, Steckhan N, Eyben F, Schuller BW, Arnrich B. Voice analysis for neurological disorder recognition–a systematic review and perspective on emerging trends. *Front Digit Health*. 2022;4:842301. doi:10.3389/fdgth.2022.842301
25. Perlis RH, Iosifescu DV, Castro VM, et al. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. *Psychol Med*. 2012;42:41-50. doi:10.1017/s0033291711000997
26. McCoy TH, Castro VM, Rosenfield HR, Cagan A, Kohane IS, Perlis RH. A clinical perspective on the relevance of research domain criteria in electronic health records. *Am J Psychiat*. 2015;172:316-320. doi:10.1176/appi.ajp.2014.14091177
27. Perlis RH, Fava M, McCoy TH. Can electronic health records revive central nervous system clinical trials? *Mol Psychiatr*. 2019;24:1096-1098. doi:10.1038/s41380-018-0278-z
28. McCoy TH, Yu S, Hart KL, et al. High throughput phenotyping for dimensional psychopathology in electronic health records. *Biol Psychiat*. 2018;83:997-1004. doi:10.1016/j.biopsych.2018.01.011
29. McCoy TH, Castro VM, Roberson AM, Snapper LA, Perlis RH. Improving prediction of suicide and accidental death after discharge from general hospitals with natural language processing. *Jama Psychiat*. 2016;73:1064. doi:10.1001/jamapsychiatry.2016.2172
30. Barroilhet SA, Pellegrini AM, McCoy TH, Perlis RH. Characterizing DSM-5 and ICD-11 personality disorder features in psychiatric inpatients at scale using electronic health records. *Psychol Med*. 2020;50:2221-2229. doi:10.1017/s0033291719002320
31. Hart KL, Pellegrini AM, Forester BP, et al. Distribution of agitation and related symptoms among hospitalized patients using a scalable natural language processing method. *Gen Hosp Psychiat*. 2021;68:46-51. doi:10.1016/j.genhosppsych.2020.11.003
32. McCoy TH, Castro VM, Rosenfield HR, Cagan A, Kohane IS, Perlis RH. A clinical perspective on the relevance of research domain

criteria in electronic health records. *Am J Psychiat*. 2015;172:316-320. doi:10.1176/appi.ajp.2014.14091177
33. Filannino M, Stubbs A, Uzuner Ö. Corrigendum to "Symptom severity prediction from neuropsychiatric clinical records: overview of 2016 CEGS N-GRID shared tasks Track 2" [J Biomed Inform. 2017 Nov;75S:s62-S70]. *J Biomed Inform*. 2018;85:204. doi:10.1016/j.jbi.2018.08.015
34. McCoy TH, Pellegrini AM, Perlis RH. Differences among Research Domain Criteria score trajectories by Diagnostic and Statistical Manual categorical diagnosis during inpatient hospitalization. *PLoS One*. 2020;15:e0237698. doi:10.1371/journal.pone.0237698
35. McCoy TH, Han L, Pellegrini AM, Tanzi RE, Berretta S, Perlis RH. Stratifying risk for dementia onset using large-scale electronic health record data: a retrospective cohort study. *Alzheimer's Dementia*. 2019;16:531-540. doi:10.1016/j.jalz.2019.09.084
36. McCoy TH, Pellegrini AM, Perlis RH. Research Domain Criteria scores estimated through natural language processing are associated with risk for suicide and accidental death. *Depress Anxiety*. 2019;36:392-399. doi:10.1002/da.22882
37. Braak H, Braak E. Demonstration of amyloid deposits and neurofibrillary changes in whole brain sections. *Brain Pathol*. 1991;1:213-216. doi:10.1111/j.1750-3639.1991.tb00661.x
38. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's. *Dementia*. 2011;7:263-269. doi:10.1016/j.jalz.2011.03.005
39. Dubois B, Hampel H, Feldman HH, et al. Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. *Alzheimers Dement*. 2016;12(3):292-323. doi:10.1016/j.jalz.2016.02.002
40. Jack CR, Bennett DA, Blennow K, et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease. Alzheimer's. *Dementia*. 2018;14:535-562. doi:10.1016/j.jalz.2018.02.018
41. Sullivan K, Pantazopoulos H, Liebson E, et al. What can we learn about brain donors? *Handb Clin Neurol* 2018;150:181-196. doi:10.1016/b978-0-444-63639-3.00014-1
42. Hart KL, Perlis RH, McCoy TH. Mapping of transdiagnostic neuropsychiatric phenotypes across patients in two general hospitals. *J Acad Consult Psychiatry*. 2021;62:430-439. doi:10.1016/j.jaclp.2021.01.002
43. Nelson PT, Alafuzoff I, Bigio EH, et al. Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature. *J Neuropathol Exp Neurol*. 2012;71:362-381. doi:10.1097/nen.0b013e31825018f7
44. Nelson PT, Jicha GA, Schmitt FA, et al. Clinicopathologic correlations in a large Alzheimer disease center autopsy cohort: neuritic plaques and neurofibrillary tangles "do count" when staging disease severity. *J Neuropathol Exp Neurol*. 2007;66:1136-1146. doi:10.1097/nen.0b013e31815c5efb
45. Hyman BT, Phelps CH, Beach TG, et al. National institute on aging–Alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's Dementia*. 2012;8:1-13. doi:10.1016/j.jalz.2011.10.007
46. Ross CA, Margolis RL. Research domain criteria: strengths, weaknesses, and potential alternatives for future psychiatric research. *Mol Neuropsychiatry*. 2019;5:218-236. doi:10.1159/000501797

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Vogelgsang J, Dan S, Lally AP, et al. Dimensional clinical phenotyping using *post-mortem* brain donor medical records: *post-mortem* RDoC profiling is associated with Alzheimer's disease neuropathology. *Alzheimer's Dement*. 2023;15:e12464. https://doi.org/10.1002/dad2.12464