

Natural Language Processing for Identification of Negative Symptoms in Patients with and without Cognitive Impairment Associated with Schizophrenia (CIAS)

SA71

Theresa Cassidy¹, Ashley Wu², Benjamin Fell³, Monika Frysz⁴, Georgina Ireland³, Suzanne St.Rose², Ceyda Uysal³, Rashmi Patel⁵

1. Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT, USA; 2. Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany; 3. Akrivia Health, Department of Research, Oxford, UK; 4. Boehringer Ingelheim, Ltd., Real World Evidence CoE, Bracknell, UK; 5. Department of Psychiatry, University of Cambridge, Cambridge, UK

Introduction

- 1

CONTEXT

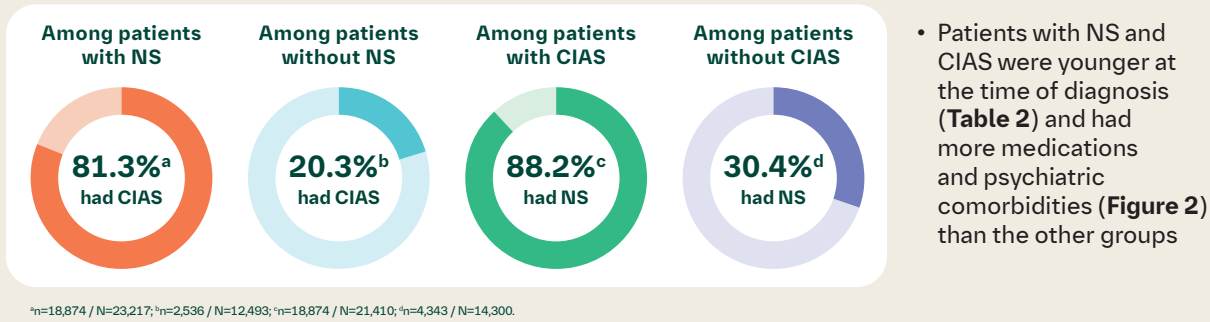
- CI in patients with schizophrenia is associated with greater disease severity, functional impairment and HCRU relative to patients without CIAS^{1,2}
 - CI is characterised by deficits in cognitive functions (e.g. speed of processing, verbal learning, working memory, reasoning/problem solving and social cognition)¹
 - An NLP method was developed by Mascio et al. to aid identification of cognitive problems from the EHRs of patients with schizophrenia² and replicated on data from 12 NHS Trusts in the Akrivia Health dataset³
- 2

UNMET NEED

- Although previously considered for inclusion in diagnosis reference manuals (e.g. DSM-5),⁴ there remains no standardised diagnostic code for CIAS

Key Conclusion

Figure 1. CIAS and NS commonly occurred together



Aims

The aim of this study was to use NLP to identify patients with CIAS and NS in order to compare demographic and clinical characteristics and understand how CIAS/NS status impacts medication use and prevalence of comorbidities

Methods

Study design

- Non-interventional study (1 Jan 2005 to 31 Dec 2023)
- Data source: Akrivia Health dataset

- Real-world data from mental healthcare services in 12 NHS Trusts in England and Wales
 - De-identified and anonymised data from 4.6 million patients
 - CIAS status determined from clinical notes' free text by NLP

- Key inclusion criteria

- ≥18 years of age at schizophrenia diagnosis
 - First schizophrenia diagnosis (according to ICD-10 code F20) within the study period
 - Clinical notes within the study period
- Key exclusion criteria

- Presence of dementia, mild CI or intellectual disability
 - First CI mention >3 months post-diagnosis
 - First diagnosis of schizophrenia outside the study period
 - Non-binary gender or gender information missing

- Outcomes
- Common NS keywords

Patient characteristics

Medication use

Comorbidities

- Data analysis

- 35 clinician-identified keywords indicating NS were used to determine patient NS status. Some examples include:

Social withdrawal

Anhedonia

Speech limited

Flat affect

Poor eye contact

Emotionally subdued

Needs prompting

Indifference

Loss of interest

Impaired social interaction

Low motivation
 - For the full list of clinician-identified keywords, see **Supplementary Table 1** on the conference app
 - NS keyword prevalence is reported descriptively
 - The NS domain of the PANSS was used as a reference framework
 - Statistical hypothesis testing was conducted using the Kruskal-Wallis test for continuous variables and Chi-squared test for categorical variables to evaluate differences by schizophrenia subgroups
 - Effect sizes were estimated using Eta-squared and Cramer's V

Results

Study population

- Of 35,710 patients identified with schizophrenia, 23,217 (65.0%) were classed as having NS, 18,874 (81.3%) of whom had CIAS (**Figure 1**)

Common NS keywords

Table 1. The 10 most common NS key words in patients with schizophrenia, ordered by prevalence in patients with CIAS	With CIAS	Without CIAS
Withdrawn behaviour	13,889 (64.9%)	2,144 (15.0%)
Social interaction finding	7,676 (35.9%)	467 (3.3%)
Flat affect	7,630 (35.6%)	459 (3.2%)
Impaired concentration	7,578 (35.4%)	0 (0%)
Poor eye contact	7,464 (34.9%)	34 (0.2%)
Low motivation	7,068 (33.0%)	637 (4.5%)
Impaired social interaction	6,648 (31.1%)	304 (2.1%)
Self-neglect	5,199 (24.3%)	577 (4.0%)
Finding related to ability to perform personal hygiene activity	4,517 (21.1%)	191 (1.3%)
Finding of general energy	4,073 (19.0%)	376 (2.6%)

Patient characteristics

- Patients with both NS and CIAS were more likely to be male than female and had a lower mean age at diagnosis compared with other NS/CIAS groups (**Table 2**)

Table 2. Patient characteristics by NS and CIAS status							
		Without NS and without CIAS (n=9,957)	Without NS and with CIAS (n=2,536)	With NS and without CIAS (n=4,343)	With NS and with CIAS (n=18,874)	P-value	Effect size
Age at diagnosis, years	Mean (SD)	46.4 (15.9)	45.4 (16.6)	45.0 (15.2)	41.3 (14.5)	<0.001	η², 0.02 (small)
	Median (IQR)	45 (34–57)	44 (32–56)	43 (33–55)	40 (30–51)		
Sex, n (%)	Female	3,566 (35.8)	1,006 (39.7)	1,500 (34.5)	7,127 (37.8)	<0.001	Cramer's V, 0.03 (minimal)
	Male	6,391 (64.2)	1,530 (60.3)	2,843 (65.5)	11,747 (62.2)		

Medication use

- Patients with NS and CIAS had more medications mentioned in the 5 years following diagnosis than patients without NS and CIAS
- The median (IQR) number of medications in patients without NS/without CIAS, without NS/with CIAS, with NS/without CIAS and with NS/with CIAS was 1 (0–2), 2 (1–4), 2 (1–4) and 5 (2–11), respectively (p<0.001; effect size η², 0.13 [medium])
- A higher proportion of patients with NS and CIAS had second-generation antipsychotics mentioned (78.7%) than patients with NS without CIAS (55.6%) and patients without NS with CIAS (53.4%)
- Benzodiazepine derivative sedatives/hypnotics were also more common in patients with NS and CIAS (45.4%) compared with patients without NS with CIAS (20.5%) and patients with NS without CIAS (17.1%)

Comorbidities

- In patients with both NS and CIAS, the prevalence of psychiatric (**Figure 2**) and physical (**Figure 3**) comorbidities was generally higher than in the other groups

Figure 2. Frequency of psychiatric comorbidities by NS and CIAS status

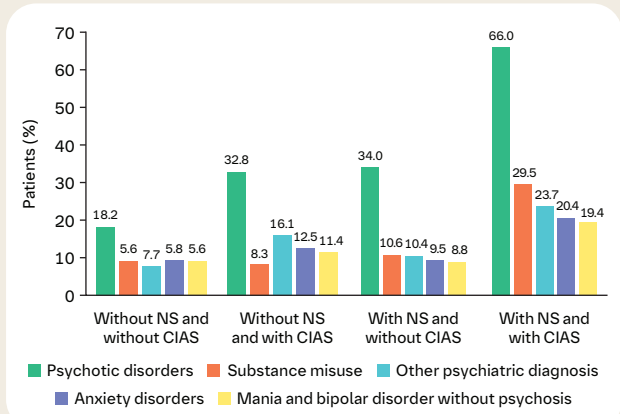
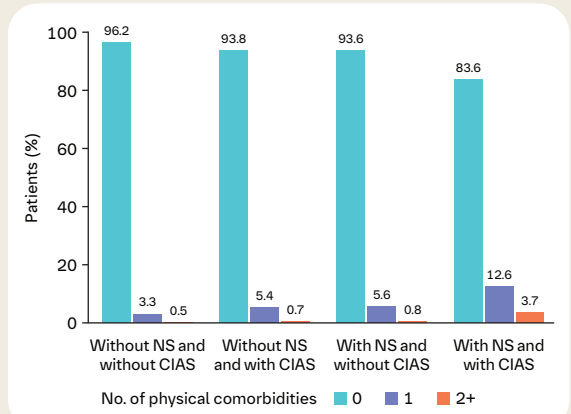


Figure 3. Frequency of physical comorbidities by NS and CIAS status



Abbreviations

CI, cognitive impairment; DSM-5, Diagnostic and Statistical Manual of Mental Disorders, 5th Edition; EHR, electronic health record; HCRU, healthcare resource utilisation; ICD-10, International Classification of Diseases, 10th Revision; IQR, interquartile range; NHS, National Health Service; NLP, natural language processing; NS, negative symptoms; PANSS, Positive and Negative Syndrome Scale; SD, standard deviation.

References

1. Green MF, et al. *World Psychiatry* 2019;18:146–61. 2. Mascio A, et al. *Front Digit Health* 2021;3:711941. 3. Todorovic A, et al. *BMJ Open* 2024;14:e088166. 4. McCutcheon RA, et al. *Mol Psychiatry* 2023;28:1902–18.

Disclosures

TC is an employee of Boehringer Ingelheim Pharmaceuticals, Inc. **AW** and **SSR** are employees of Boehringer Ingelheim International GmbH. **MF** is an employee of Boehringer Ingelheim, Ltd. **BF**, **GI** and **CU** are employees of Akrivia Health. **RP** has received grant funding from the National Institute for Health and Care Research (NIHR301690), the Medical Research Council (MR/S003118/1), the Academy of Medical Sciences (SGL015/1020) and Janssen, has participated in Scientific Advisory Boards for Boehringer Ingelheim and Teva, and has received consulting fees from Holmusk, Akrivia Health, Columbia Data Analytics, Clinilabs, Social Finance, Boehringer Ingelheim, Bristol Myers Squibb, Supernus, Teva and Otsuka.

Acknowledgements

The authors meet the criteria for authorship as recommended by the International Committee of Medical Journal Editors. This study was funded by Boehringer Ingelheim. Writing, editorial support, and formatting assistance were provided by Heather St Michael, BSc, of Avalere Health Global Limited, funded by Boehringer Ingelheim International GmbH.

Additional Conclusion

- Applying NLP to EHRs may enable earlier identification of NS in patients with CIAS and ensure the unmet needs in this patient population are duly recognised

