The Benefits of a Complex Identification and Stratification Model for Depression: Insights from a Large Claims Analysis Study

Amber Shergill MPH[1], Andrey A. Popkov MS[1], Tyson S. Barrett PhD[1]
Jason Hohl MBA[1], Susan L. Deakin DO[1], Melissa Perry MD[1]

Introduction

Accurately evaluating depression severity is imperative for healthcare providers and insurers to determine appropriate treatment approaches using the stepped care approach. Reliably stratifying patients into mild, moderate and severe categories based on validated cut-off scores can help clinicians customize treatment plans and improve outcomes A challenge in accurately evaluating depression severity is the fragmentation of health data across multiple systems and incomplete records.[1,2] Variables needed to score severity on multidimensional scales are often spread across disparate electronic medical records, claims databases, and provider networks.[3,4] This fragmentation creates barriers to properly identifying and classifying patient populations based on depression severity in order to target appropriate interventions.[5,6] Advanced analytics approaches leveraging machine learning methods have shown promise in integrating heterogeneous health data sources to enable more accurate phenotyping and risk stratification for conditions like depression.[7] Given the need for accurate evaluation of depression severity, this study aims to investigate the utility of an analytics-enabled identification and stratification (IDS) framework by examining its application to claims data. We utilize a variety of tools available to us as an insurance company to identify members with depression and stratify them into four levels of severity.

Objectives

To investigate an analytics-enabled identification and stratification (IDS) framework for evaluating depression, and its severity, using health insurance claims data.

Methods

This is a retrospective analysis of 2022 claims and electronic health record data from Highmark Health and affiliated insurers. Members aged 18+ with Highmark coverage and a healthcare encounter were included. The depression IDS framework used diagnosis codes, pharmacy claims, and Optum, Milliman, and John Hopkins analytics software data. Alignment of identification and stratification criteria was evaluated.

Results

The IDS framework identified 762,753 members with depression (18.8% of the population). The identification rules revealed variability in prevalence, with 2% identified by PHQ-9 scores, 11% by diagnoses, 9% by treatment groups, and 11% by adjusted clinical groups. The framework identified 306,394 more members (7.6% of the population) compared to using diagnoses alone. The IDS rules escalated 46% of mild and 19% of moderate cases to higher severity compared to single parameter assessments. Expenses for severe depression were 159% higher than for minimal severity.

Characteristic	Mean/Count (SD/%)	Prevalence
	n = 773,166	
Sex		
Female	522,295 (67.6%)	25.3%
Male	250,677 (32.4%)	12.8%
Other/Unknown	194 (0%)	20.4%
Age		
18-30 years old	135,749 (17.6%)	14.9%
31-40 years old	115,771 (15%)	17.1%
41-50 years old	128,154 (16.6%)	19.2%
51-60 years old	142,881 (18.5%)	19.2%
61-70 years old	129,155 (16.7%)	22.2%
70 years old or older	121,456 (15.7%)	27.5%
Insurance Type		
ACA	70,335 (9.1%)	24.1%
Commercial	485,052 (62.7%)	18.4%
Medicare Advantage	91,513 (11.8%)	32.5%
Medicaid	28,068 (3.6%)	31.4%
Other	98,198 (12.7%)	13.7%

Table 1. Descriptive statistics of the sample (n = 765,688) and prevalence of depression identification by each characteristic.

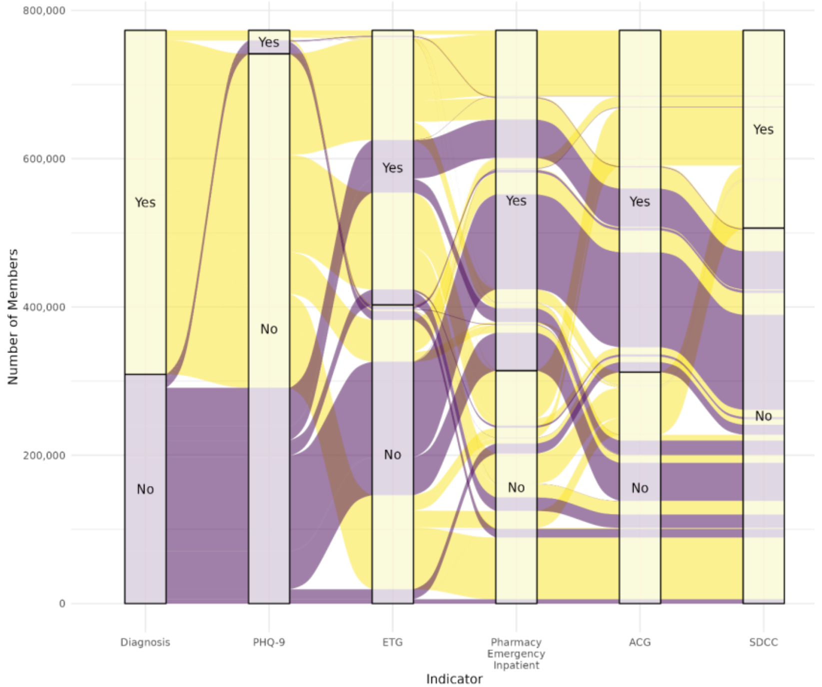

Figure 1. Alluvial plot showing the agreement between each depression indicator.

Discussion The operationalization of this identification and stratification model across large populations supplies a mechanism to connect people exhibiting mild depression symptoms with low-cost digital self-care tools, while reserving expensive high-touch services for those most acutely ill. This targeted approach not only improves individual patient care but also optimizes the allocation of healthcare resources, potentially leading to more sustainable healthcare systems.

Indicator	Unique Contribution	Percent of Sample
ACG	13,385	1.7
Diagnosis	82,896	10.7
PHQ-9	17,433	2.3
Pharmacy	50,844	6.6
Emergency and Inpatient	18,059	2.3
SDCC	33	< .1
2 or More Indicators	590,351	76.4

Table 2. Unique contribution of each indicator and the identification by 2 or more indicators.

Conclusions

The analytics-enabled IDS framework demonstrates utility in identifying members with depression by linking fragmented data sources. Aligning multiple parameters provides a more nuanced severity evaluation compared to individual data elements. Enhanced phenotyping enables the targeting of cost-effective digital self-care tools to milder cases while reserving higher cost interventions for the most severely ill, potentially reducing overall costs while maintaining health outcomes. Implementation of this integrative platform can help focus efforts on those with the highest need and bridge the gap in treating depression.

Acknowledgements and Permissions

Author Contributions: Mr. Popkov, Mr. Hohl, and Dr Barrett had full access to the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Price, N. (2017). Risk and resilience in health data infrastructure. *Social Science Research Network*. https://doi.org/10.2139/ssrn.2928997 2. West, S. L., Johnson, W. G., Visscher, W., Kluckman, M., Qin, Y., & Larsen, A. (2014). The challenges of linking health insurer claims with electronic medical records. *Health Informatics Journal*, *20*(1), 22–34. https://doi.org/10.1177/1460458213476506 3. Iyer, S. P., Spaeth-Rublee, B., & Pincus, H. A. (2016). Challenges in the operationalization of Mental Health quality Measures: An Assessment of Alternatives. *Psychiatric Services*, *67*(10), 1057–1059. https://doi.org/10.1176/appi.ps.201600198 4. Hanchate AD, Ash AS, Borzecki A, Abdulkerim H, Stolzmann KL, Rosen AK, Fink AS, Pugh MJ, Shokeen P, Shwartz M. How pooling fragmented healthcare encounter data affects hospital profiling. *Am J Manag Care*. 2015 Feb;21(2):129-38. PMID: 25880362 5. Amirian, P., Lang, T., & Van Loggerenberg, F. (2017). Big data in healthcare. *In Springer Briefs in pharmaceutical science & drug development*. https://doi.org/10.1007/978-3-319-62990-2 6. Kilbourne, A. M., Keyser, D. J., & Pincus, H. A. (2010). Challenges and opportunities in measuring the quality of mental health care. *The Canadian Journal of Psychiatry*, *55*(9), 549–557. https://doi.org/10.1177/070674371005500903 7. Nickson, D., Meyer, C., Walasek, L. et al. Prediction and diagnosis of depression using machine learning with electronic health records data: a systematic review. *BMC Med Inform Decis Mak* 23, 271 (2023). https://doi.org/10.1186/s12911-023-02341-x