Subgroup Identification with Bayesian Nonparametric Models for individuals with schizophrenia who are at risk for relapse

Authors: Khan, Anzalee[1,3]; Lindenmayer, JP[1,2,3]; Insel, Beverly[1]; Mena, Evan[1,4]; Seddo, Mary[1]; Parker, Benedicto[1,3]; Parak, Mohan[1,3]
Affiliations: 1) Manhattan Psychiatric Center, New York, NY 2) New York University, Department of Psychiatry, New York, NY 3) Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY 4) St George's University, Grenada, WI

METHODOLOGIAL QUESTION

Many individuals with schizophrenia relapse within 1 year following discharge from inpatient hospitalization. Although prior research has identified several risk factors for relapse among individuals with schizophrenia, little is known about how these factors may interact to modify relapse risks. We conducted an exploratory analysis aimed to obtain an empirically based set of interactions of factors related to relapse rates of individuals with schizophrenia. Unlike logistic regression analyses that include only main effects or interaction terms specified a priori, these analyses allow for the investigation of more complex interactions.

Can Bayesian nonparametric models be used to identify subgroups with high or low rates of relapse through the presentation of decision trees of complex interactions in a way that is likely to be readily understandable by clinicians and policy makers?

AIMS

The aim of this study is to identify clinical and demographic variables that form a subgroup of individuals who are likely at risk for relapse

DEFINITION OF RELAPSE

To determine factors predicting relapse in subjects with chronic schizophrenia, a Cox regression survival analysis with backward elimination modeling was performed. Time to first relapse within the first two years of follow-up was included as the dependent variable, and it was defined as:

- Arrest/Incarceration
- Psychiatric Hospitalization
- Suicide/Self-Harm
- Discontinuation of Antipsychotic medication that leads to hospitalization
- Treatment supplementation with another antipsychotic that leads to hospitalization
- Increase in the level of psychiatric services in order to prevent imminent psychiatric hospitalization

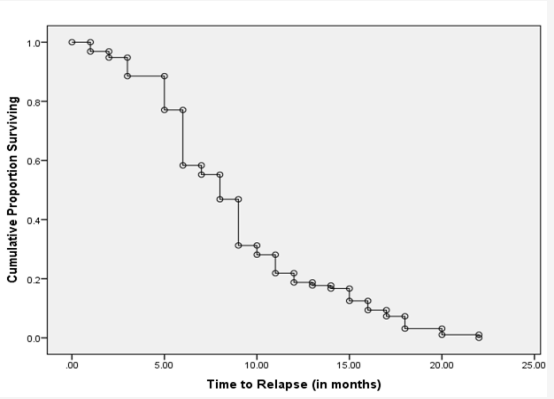
METHODS

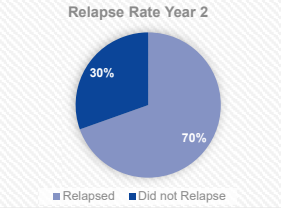
Sample: We assessed 140 inpatients discharged from a psychiatric facility in New York, NY and followed for two years
Analysis: Data mining techniques (Bayesian Dirichlet Equivalent (BDE)) were used to reanalyze data from a retrospective study by randomly splitting the data into two samples (primary (n = 70) and replication (n = 70)), and developing a decision tree for the primary sample using recursive partitioning. We then tested whether the subgroups developed within the primary sample were associated with increased relapse risk in the replication sample
Predictors: Analyses focused on predicting relapse with the following candidate risk factors: race, treatment with a Long Acting Injectables, treatment with clozapine, sex, education, MCCB domain scores, age, substance use disorder diagnosis, comorbid psychiatric disorder, prior hospitalizations (< > 5), length of recent hospitalization, Marder PANSS factors, PANSS G12 Lack of judgment and Insight, and baseline medical comorbidity. A cut-off of p<.05 was used as an indicator of statistical significance.

Baseline Sample Characteristics

	N = 138
	Mean (SD)
Age (years)	42.12 (7.23)
Education (years)	11.36 (3.68)
Age of onset (years)	20.11 (3.22)
Substance abuse prior to admission (years)	10.36 (4.68)
Number of previous hospitalization	9.16 (6.43)
Length of stay (prior to discharge) (months)	12.01 (6.23)
	%
Gender	
Male	87.68%
Female	12.32%
Antipsychotic Treatment	
Oral antipsychotics[a]	58.70%
Intramuscular depot[b]	41.30%
Ethnicity	
African American	55.07%
Asian	3.62%
Caucasian	21.01%
Hispanic	20.29%
Baseline Symptoms	Mean (SD)
PANSS Positive Subscale	18.03 (4.69)
PANSS Negative Subscale	21.36 (5.67)
PANSS Total Score	78.22 (12.48)
Clinical Global Impression Scale - Severity	4.1 (0.98)
Cognitive Symptoms	Mean (SE)
Speed of Processing	14.15 (1.01)
Attention/Vigilance	18.53 (1.56)
Working Memory	10.69 (1.35)
Verbal Learning	26.45 (0.49)
Visual Learning	23.48 (0.79)
Reasoning & Problem Solving	29.35 (0.62)
Global Composite T Score	18.69 (36.44)
Functional Outcomes	
Personal and Social Performance (PSP) Global	70.24 (10.22)

Relapse Rate (2 years following discharge)





	Mean				Median			
			95% Confidence Interval				95% Confidence Interval	
	Est	Std. Error	Lower Bound	Upper Bound	Est	Std. Error	Lower Bound	Upper Bound
	8.781	.481	7.839	9.724	8.000	.425	7.167	8.833

30.43% (n=42) of the population had multiple relapses (defined as >2 relapses in 2 years)

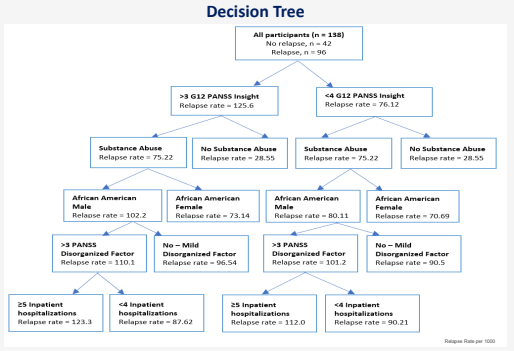
Bayesian Nonparametric Model

The analysis produced a decision tree with subgroups at differing levels of risk for relapse. These were identified by a combination of factors:

- >3 on PANSS G12 Lack of Judgment and Insight,
- co-occurring substance use,
- African American race, male,
- PANSS Disorganization factor score, and
- ≥ 5 previous inpatient hospitalizations

The groups developed as part of the decision tree accurately discriminated between those with and without relapse in the replication sample:

- Item G12 on the PANSS, PANSS Disorganized Factor, substance use, and being African American were the single strongest indicators of relapse (OR = 2.1 (95%CI 1.5, 3.3)
- A series of comparisons were conducted in the replication sample to examine the extent to which more traditional analytic techniques would yield similar results. In all cases, chi-square test showed the groups produced by the initial exploratory data mining analyses significantly discriminated those who relapsed from those who did not.



The data mining analysis identified poor judgement and insight as the strongest single risk factor for relapse

CONCLUSIONS

This study uses an empirically derived decision tree to identify subgroups by relapse rates. Results indicate that the association between a known risk factor and relapse is not independent of another risk factor. Findings suggest that the identification of individuals at increased risk for relapse is improved through the examination of higher order interactions between risk factors. This significant variability in level of risk depending on the interaction of a group could facilitate the study of how these risk factors interact and combine to influence relapse. Clinicians treating these subgroups should be aware of these risks and consider targeting treatment strategies to address these combined factors.

Financial Disclosures The authors report no conflicts of interest for this work. Contact: anzalee.khan@nki.rfmh.org