



# HHS Public Access

Author manuscript

*J Ment Health Policy Econ.* Author manuscript; available in PMC 2024 June 01.

Published in final edited form as:

*J Ment Health Policy Econ.* 2023 June 01; 26(2): 63–76.

## Applying Machine Learning to Human Resources Data: Predicting Job Turnover among Community Mental Health Center Employees

**Sadaaki Fukui, Ph.D., MSW,**

Associate Professor, Indiana University School of Social Work, 902 West New York Street,  
Indianapolis, IN 46202-5156 USA

**Wei Wu, PhD,**

Associate Professor, Psychology, Indiana University-Purdue University, Indianapolis, 402 N.  
Blackford St., Indianapolis, IN 46202, USA

**Jaime Greenfield, MS,**

Vice President of Operations, Places for People, 1001 Lynch Street, Saint Louis, MO 63118, USA

**Michelle P. Salyers, Ph.D,**

Professor, Indiana University-Purdue University Indianapolis, Department of Psychology, 402 N.  
Blackford St., LD124 Indianapolis, IN 46202-3217 USA

**Gary Morse, Ph.D,**

Former Vice President Research and Development, Places for People, 4130 Lindell Blvd. St.  
Louis, MO 63108, USA

**Jennifer Garabrant, BSW,**

Program Manager, ACT Center of Indiana, Indiana University-Purdue University Indianapolis,  
Department of Psychology, 402 N. Blackford, LD 120B, Indianapolis, IN 46202, USA

**Emily Bass, BA,**

Graduate Student, Psychology, Indiana University-Purdue University, Indianapolis, 402 N.  
Blackford St., Indianapolis, IN 46202, USA

**Eric Kyere, Ph.D,**

Assistant Professor, Indiana University School of Social Work, 902 West New York Street,  
Indianapolis, IN 46202-5156 USA

**Nathaniel Dell, MA, MSW, LCSW**

Vice President of Knowledge Translation and Impact, Places for People, 4130 Lindell Blvd. St.  
Louis, MO 63108, USA

---

Address Correspondence to: Sadaaki Fukui, Associate Professor, Indiana University School of Social Work, 902 West New York Street, Indianapolis, IN 46202-5156 USA, TEL: 317-278-1767; sadafuku@iu.edu.

This is the author's manuscript of the article accepted as: Fukui, S., Wu, W., Greenfield, J., Salyers, MP., Morse, G., Garabrant, J., Bass, E., Kyere, E., & Dell, N. (accepted). Applying Machine Learning to Human Resources Data: Predicting Job Turnover among Community Mental Health Center Employees. *The Journal of Mental Health Policy and Economics*.

Disclosures:

The content is solely the responsibility of the authors and does not represent the official views of NIH.

## Abstract

**Background:** Human resources (HR) departments collect extensive employee data that can be useful for predicting turnover. Yet, these data are not often used to address turnover due to the complex nature of recorded data forms.

**Aims of the Study:** The goal of the current study was to predict community mental health center employees' turnover by applying machine learning (ML) methods to HR data and to evaluate the feasibility of the ML approaches.

**Methods:** Historical HR data were obtained from two community mental health centers, and ML approaches with random forest and Lasso regression as training models were applied.

**Results:** The results suggested a good level of predictive accuracy for turnover, particularly with the random forest model (e.g., Area Under the Curve was above .8) compared to the Lasso regression model overall. The study also found that the ML methods could identify several important predictors (e.g., past work years, wage, work hours, age, job position, training hours, and marital status) for turnover using historical HR data. The HR data extraction processes for ML applications were also evaluated as feasible.

**Discussion:** The current study confirmed the feasibility of ML approaches for predicting individual employees' turnover probabilities by using HR data the organizations had already collected in their routine organizational management practice. The developed approaches can be used to identify employees who are at high risk for turnover. Because our primary purpose was to apply ML methods to estimate an individual employee's turnover probability given their available HR data (rather than determining generalizable predictors at the wider population level), our findings are limited or restricted to the specific organizations under the study. As ML applications are accumulated across organizations, it may be expected that some findings might be more generalizable across different organizations while others may be more organization-specific (idiographic).

## Keywords

machine learning; human resources data; employee turnover; community mental health center; data-driven management strategies; job retention

---

Employee turnover continues to be a critical problem for many community mental health workers, with rates of turnover ranging from 25% to 60% annually (1–3). Vacancies in positions often remain unfilled for long periods (4). Indeed, some data indicate that over 70% of counties in the U.S. report a severe shortage of mental health professionals (5, 6). High rates of voluntary turnover negatively affect both the organization (e.g., extra cost for new hires and training, reduced productivity) and individual employees (e.g., diminished well-being, reduced financial stability) (1, 3, 7–9), subsequently decreasing the quality of care provided for clients (3, 8, 10–12). Thus, it is imperative for employee well-being, quality of care, and cost considerations to develop methods to identify and prevent avoidable turnover among community mental health employees.

Turnover research in mental health is still limited and generally relies on cross-sectional surveys with few variables and small convenience samples. Most existing research has

identified factors correlated with turnover *intentions*, a proxy for actual turnover (13). Some potential turnover factors suggested in mental health include insufficient salary, increased burnout, decreased job satisfaction, the lack of organizational support, and lack of professional development opportunities (14). Additional negative factors include job demands (15), emotional labor, low organizational trust (14), workplace interpersonal conflict, and heavy workload (16). Conversely, positive factors that may reduce turnover include positive workplace climates, fair performance appraisal, job autonomy, and workplace psychological safety (11).

Unfortunately, turnover predictors suggested in the literature may not always be applicable to a specific organization, given differing characteristics (e.g., demographic diversity, organizational size, different management systems, and clientele). Identifying specific variables that require attention in a particular organizational setting is challenging. Collecting more data from employees (who are already often overwhelmed with existing documentation requirements) might not be well tolerated, especially when there is a lack of clarity about the salient predictors. Further, administrators and evaluators may not be able to collect information from those disengaged in their work (especially those intending to leave). It may also be difficult, if not impossible, to identify an individual at high turnover risk through anonymous survey methods.

Human Resources (HR) departments typically collect extensive employee data, including demographics, the type and nature of the job, withdrawal behaviors (e.g., absence), and job performance, which are presented in turnover theories (17). In mental health organizations in particular, other administrative data such as the client populations being served (e.g., client caseload, client symptom severity, age, missed appointments) and work contexts (e.g., individual vs. team, work department) may also be available. Such a wealth of information could help us understand employee turnover at the organizational level. For example, Rombaut and Guerry (18) found that available HR data (e.g., demographics; work-specific factors, such as seniority and salary) could result in reliable turnover predictions in a private company in Belgium. However, HR data are often collected in different data management systems, with sparse and different formats, which are difficult to handle with traditional data analysis methods. As a result, these data are not often analyzed to address turnover in mental health organizations.

Although the application of machine learning (ML) methods in practice is still limited, applying these analytics to HR data for predicting employee turnover is emerging as a promising approach (19). ML, which can handle extensive heterogeneous data, involves a set of computational strategies or algorithms to recognize and discover data patterns systematically. ML creates automated algorithms to learn from data and make data-driven predictions or decisions (20). ML algorithms have gained popularity in HR analytics as a data-driven approach when predicting turnover (21–23). For example, Ribes et al. (24) used ML to identify meaningful predictors for employee turnover, such as business units, employee behavior, and performance. Sajjadi et al. (25) applied ML to job application data to predict job performance and turnover among applicants for public school teaching positions. Quinn et al. (26) also applied ML to HR data for predicting turnover among child welfare workers. Additionally, Kuncel et al. (27) conducted a meta-analysis comparing

the predictive power of mechanical methods (i.e., applying an algorithm or formula) versus holistic methods (i.e., clinical expert judgment, intuitive, subjective) and found that mechanical methods outperformed holistic methods in areas such as predicting job performance.

ML algorithms applied to HR data could identify organization-based, targeted variables when addressing turnover in a specific organizational context. Further, the algorithms developed in a particular organizational context may be adapted to other organizations, tailoring the turnover predictions and continuing to improve precision in both unique (local) and more generalizable (global) contexts. Successful ML applications will suggest data-driven, evidence-based HR management strategies for organizations to prevent employee turnover. However, as far as the authors are aware, ML methods have not been applied to HR data to predict turnover among community mental health organization employees.

The goals of the current study are three-fold: 1) applying ML methods to historical HR data for predicting an employee's turnover probability within the following 12 months; 2) exploring predicting variables; and 3) evaluating the feasibility of ML applications to HR data at community mental health organizations. Because HR data management systems or available data are different across different centers, we developed ML predictive models for one center first (the main evaluation center) and used the center to present the primary ML application process. We then applied the same method to another center (the application center) to examine the applicability of the ML approaches in another setting.

## Methods

### Study Setting

Historical HR data were obtained from a community mental health center in an urban midwestern city (our “main evaluation center”). The organization had approximately 300 employees and provided services that included case management, home-based, and school-based services, supported employment, medication management, and outpatient individual and group therapy. The HR and administrative data extraction involved six staff from the HR, Operations, and Clinical Administrations departments at the center using five different HR data management systems. The data were de-identified (with research IDs), cleaned, and processed for the ML predictive model development and application by the research team. The study was approved by the [university] Institutional Review Board, and waiver of informed consent was applied to the data already being collected by the organizations.

### Data Preparation and Processing

The HR data (dating back to 2011) contained 751 employees' records, including those who had already left the job (leavers) and those who stayed (stayers) at the time of data extraction (January 2021). Employee IDs were replaced with research IDs which were used to merge the datasets for ML. We excluded the following cases: 1) seven cases that never started the job, 2) 20 cases with missing data in their hire date, 3) ten cases hired in 2021, given the lack of sufficient historical information for ML, and 4) 60 involuntary leavers due to reasons such as unsatisfactory job performance, ended funding, or job elimination. Thus, 654 cases

were used in the analysis. Among them, 336 cases (51.4%) left the organization by the end of 2020 (leaver), and 318 (48.6%) stayed (stayer).

We extracted both categorical and continuous predicting variables (often called “features” in the ML literature) from the existing HR data. We will refer to these variables as predictors, a term most familiar to applied researchers in social and behavioral science. Categorical predictors included marital status, education level, exempt status, race, employee status/job type (i.e., clinical vs. nonclinical), and gender. Additionally, 175 job position titles were recorded, which were grouped into 12 categories (e.g., direct services) by combining work of similar nature and categories with small sizes. The categorical predictors were entered into the ML models using the dummy variable approach (28). The HR data also included work department, which contained 52 units. However, the information was not used in analyses due to a small number of cases for each department category (i.e., merging different departments was not desirable as each had unique job functions). Further, the HR data contained certification and position assignment history information. However, we did not include these data in the analyses because only a small number of employees’ information was available. Table 1 displays the frequency distributions of the categorical predictors by stayers and leavers for our main evaluation center (and application center as described below). The reference group used in dummy variable encoding is in italics in Table 1.

The continuous predictors were extracted from sources containing wage, job training, and client service data. They included employee age, years employed, salary (converted to an hourly wage), hours worked, job training hours, and client characteristics served by the employee (i.e., average age, proportions of groups by gender and diagnosis). Specifically, we calculated the employees’ age and the number of years they had worked in the organization by the previous year. For leavers (i.e., those who left at some point before the end of 2020), a date of the end of the previous year was used for the calculation. For instance, if an employee left in 2017, their age and number of work years by December 31<sup>st</sup> of 2016 were used. If the employee left in the same year when they were hired, then the number of work years would be 0. For stayers (i.e., those who had stayed until the end of 2020), we used Dec 31<sup>st</sup> of 2019 as the benchmark to calculate their age and work years.

Hourly wage and biweekly work hours were calculated for each employee. We created four variables for each person based on the salary and work hours data (i.e., mean and standard deviation of hourly wage over the years, mean and standard deviation of biweekly work hours over the years). To account for potential salary inflation, the salary was standardized by year. In addition, we calculated the total number of job trainings and the total number of training hours for each employee using the job training data (which had 3,620 elements and a total of 32,391 records). Because of the high redundancy between the two predictors (Pearson  $r > .9$ ), only the total number of training hours was used. Finally, we extracted deidentified client characteristics served by the employees, which contained 2,421 unique clients’ demographic information served by 398 employees between 2017 and 2019. The average client age, as well as the proportion of clients’ gender (i.e., male vs. female) and the clients’ diagnosis (i.e., schizophrenia vs. other diagnoses) were calculated. Note that we also had access to annual job performance evaluation data. However, we did not use the data in analyses because there was a large amount of missing data (> 50%), and these missing data

were mostly from the leavers. All the continuous predictors were standardized to remove the influence of their scales on the ML applications. The descriptive statistics of the continuous predictors by stayers and leavers are shown in Table 2 for our main evaluation center (and application center as described below).

### Missing Data

The organization introduced new HR data management systems around 2017, which created some missing data during the data transition. Missing data in the HR systems were filled by HR staff where possible for the current study. For example, about 500 employees' educational degrees were not recorded in the current HR data system, which were later transferred (e.g., Licensed Master Social Worker, or LMSW) and re-coded (i.e., diploma, associate, bachelor, master, doctoral degrees) by the HR staff. The remaining amount of missing data for the predictors included in the analyses ranged from 1.3% to 45.5%. The amount of missing data on categorical predictors was generally low (e.g., 1.3% for gender, 2.0% for marital status, 2.6% for race, 5.8% for employee type, 6.4% for job position). Missing data were imputed (i.e., replaced by plausible values) using the k-nearest neighbor (KNN), which has been shown as a generally effective missing data technique (29, 30). Briefly speaking, KNN imputes a missing value on a predictor by finding the complete cases close to the missing cases and aggregating those cases' observations to fill in the missing value. The imputed values for categorical data from KNN were continuous and recoded into categorical values by rounding to the nearest neighbor. The only exceptions were race and job position. The two predictors had more than two categories and a relatively larger proportion of missing data than the other categorical predictors. More importantly, the employees with missing position data were all leavers. Thus, instead of imputing the missing values, we replaced them with a new category representing missing data (MISS) (31). We included the MISS category (which may not be meaningfully interpreted) only to prevent the ML models from losing cases. For race, there were only 3 Asian or Pacific islander cases, which were also combined into the MISS category to keep these cases in the analysis.

### ML Applications

We adopted both parametric and nonparametric ML approaches in our study. Parametric approaches are more parsimonious and easier to interpret (32). However, they require distributional assumptions and correct specifications about variable relationships. Consequently, they have limited capability to account for complex nonlinear relationships in data. Nonparametric ML approaches, in comparison, are designed to overcome such limitations. However, many nonparametric approaches are black-box approaches (e.g., not clearly specifying what or how interactions were considered, directionality in variable association), resulting in challenges in interpreting the results. Accordingly, we used both types of approaches to complement each method's strengths and shortfalls. As a preliminary study, we choose one representative method from each type.

For the parametric approach, we choose Lasso regression given its close relationship with logistic regression (the most popular traditional method for classification problems). Lasso regression is a regularization technique that extends logistic regression analysis by adding a penalty term for regression coefficients in the discrepancy function for estimation (33, 34).

This penalty term will shrink the estimates for coefficients, leading to biased estimates. However, this is intentionally designed to achieve a better variance-bias tradeoff (i.e., reduce the variance of prediction at the expense of a slight increase in bias), consequently improving the generalizability of the result (35). Compared to logistic regression, Lasso regression is capable of handling more effects (e.g., interaction effects) or parameters and is often used to select predictors. Thus, one could fit more complex models than logistic regression. To establish a suitable Lasso regression model, we started with the simplest model without interaction effects and then added all possible two-way interactions. However, because the model with all two-way interactions did not lead to a better prediction, we adopted the Lasso regression model without interaction effects in the current study.

For the nonparametric approach, we used random forest, which has shown superior performance for classification problems in the existing literature (32, 36, 37). Random forest is a powerful extension to the classification/decision tree (i.e., a classic nonparametric method for classification). Briefly speaking, the classification tree predicts a response variable by successively splitting the data set into increasingly homogeneous subsets based on one predictor at a time (38). Although the classification tree accounts for nonlinear relationships, a single classification tree is prone to noise (e.g., errors, residuals) in a training set, limiting its generalizability (39, 40). Random forest solves the problem by assembling results across a large number of trees. Specifically, it generates many samples based on the original data using a resampling approach such as bootstrapping and builds a tree for each sample. Furthermore, it allows the random selection of a subset of predictors at each split when building a tree, reducing the correlations or redundancy among the trees and improving the generalizability of the result (41). More information on random forest can be found in Kuhn and Johnson (29).

Both Lasso regression and random forest involve tuning hyperparameters (i.e., their values control the learning process). For Lasso regression, we tuned the so-called regularization parameter  $\lambda$ , which can range anywhere from 0 to 1. Increase of  $\lambda$  would decrease variance but increase bias in the estimates. For random forest, several parameters such as the number of trees and the number of candidate or plausible predictors randomly selected at each split of the random forest process (denoted as *mtry*) could be tuned. We only tuned *mtry* in the current study, given that this parameter has exhibited a substantial impact on random forest results (32, 42). We set the number of trees at 500, which is more than sufficient to obtain stable results (43).

### Cross-Validation

To prevent overfitting (e.g., modeling random errors in a specific data point) and increase the generalizability of ML-trained predictive models, we used the repeated K-fold cross-validation (CV) approach to evaluate the performance of the predictive models. Based on the general practice and size of our data, we set  $K = 5$  (44). This approach splits the data randomly into five equal-sized folds ( $n \approx 130$  for each fold), and each fold was used as the testing data to evaluate the performance of the predictive model trained based on the other four folds. This process was repeated ten times (the participants in the five folds were different across the replications), resulting in 50 sets of predictions. The predictions

were then aggregated to evaluate the final performance of the predictive model according to various criteria.

### ML Evaluation Criteria for Prediction Power

Four criteria were used to evaluate the performance of the prediction models: overall prediction accuracy (i.e., the proportion of cases that are correctly classified), specificity (i.e., true negative rate), sensitivity (i.e., true positive rate), and area under the curves (AUC). AUC is a balance of accuracy and false positive rates, which provides an overall index of how well the model is classified across a series of thresholds. For both models (i.e., random forest and Lasso regression), the hyperparameter was chosen to maximize each criterion, resulting in four sets of results. An AUC value above .8 is considered good prediction power (45).

### ML Result Reporting and Interpretation

We used variable importance measures (VIMs) to assess the degree to which the predictors impact the turnover prediction. VIMs were calculated differently between the two methods. For Lasso regression, they were computed based on the absolute value of the t-test statistic for the corresponding model parameter. For random forest, a permutation-based approach (41) was used to calculate the VIM for each predictor. Specifically, the value of each predictor was permuted first, and the difference in prediction accuracy averaged across all trees before and after permutation was then recorded. This difference was normalized to produce the importance score for the predictor. Predictors that resulted in greater differences were considered more important or influential.

Scaled importance scores are often reported to facilitate interpretation and comparison, which are the original importance scores divided by the highest variable importance score (46). Scaled importance scores range between 0% and 100%, showing the importance of a predictor relative to the most important one in the data. For instance, a VIM of 50% means that the importance of the predictor was 50% of that of the most important one. However, there is no absolute criterion to evaluate the VIMs. They are also calculated differently across methods; thus, they are not directly comparable. In practice, VIMs are usually used to identify a set of most important predictors (e.g., top ten) (47). Given the relatively small number of predictors considered in the study, we report the top five predictors identified from each approach. For dummy variables, note that a MISS category (to maintain cases with missing data in the prediction model) is not included in the VIMs presentation as it cannot be meaningfully interpreted.

Although VIMs indicate the degree to which a predictor is influential, it does not inform the pattern of the target predicting relationship. To facilitate interpretation, we created partial dependence plots (PDPs) for the top five predictors from each approach. PDPs visualize the effect of each predictor on the probability of turnover averaged across the marginal distributions of all other predictors (48, 49). In other words, it shows how the average prediction of the outcome could change as each predictor changes (i.e., the marginal effect).

## Feasibility Evaluation of the ML Approaches to HR Data

To evaluate the feasibility of the ML approaches to HR data managed by a CMHC, we collected information from the HR director about perceived feasibility. The information included the required number of people for extracting data, estimated initial data extraction time (with the initial trials and errors), expected data extraction time (in the future), required skills for data extraction, difficulty/ease of the initial data extraction (1. *Fairly easy* to 5. *Very difficult*), and the predicted difficulty/ease of data extraction in the future trial (1. *Fairly easy* to 5. *Very difficult*).

## Applicability Evaluation of the ML Approaches

To evaluate the applicability of the ML approaches, we obtained historical HR data from another community mental health center (our “application center”) that mirrored data from the main evaluation center’s HR database for 894 unique employees. For the applicability evaluation for this study, we extracted only a subset of predictors (i.e., date of birth, gender, race, marital status, employee status, hire date, termination date, exempt status, scheduled work hours, and hourly pay rate). Note that we only extracted the most recent wage and work hours data for the application center (while we extracted all the historical wage and work hours data in multiple years for the main evaluation center) for the current evaluation. Accordingly, we could not calculate the mean and standard deviation of these variables nor account for inflation. We extracted the HR data in the time window between 2017 and 2020 (the center introduced a new HR data management system in 2017). The size and services of the center were similar to those of the main evaluation center, but the application center was located in a rural midwestern town. Among the employees, there were 407 (45.5%) stayers and 487 (54.5%) leavers at the time of data extraction. All the predictors were complete except for marital status (5% missing) and gender (0.4% missing). The same ML approaches used for the main evaluation center were applied. The frequency distributions for categorical predictors can be found in Table 1, with the reference group highlighted in italics. The descriptive statistics of the continuous predictors are shown in Table 2.

## Software Implementation

All analyses were implemented in R 4.1.2 (50). Specifically, the ML applications were implemented using the caret package in R (42).

## Results

### Results for the Main Evaluation Center

Table 3 shows the means and standard deviations of the AUC, sensitivity, specificity, and overall accuracy measures from the 50 sets of predictions from the repeated K-fold cross-validation for the Lasso regression and random forest approaches. Paired *t*-tests were used to test whether the means differed significantly between the two approaches. To avoid inflated type I error rates, the  $\alpha$  value for the *t*-tests was adjusted using Bonferroni’s correction for multiple outcomes (four criteria, adjusted  $\alpha = 0.05/4 = 0.0125$ ). The result suggested that random forest performed significantly better than Lasso regression for all criteria except for sensitivity: AUC [difference = .09,  $t(49) = 13.70$ ,  $p < .01$ ], specificity [difference = 0.25,

$t(49) = 20.35, p < .01$ ], and overall accuracy [difference = 0.12,  $t(49) = 18.30, p < .01$ ]. For sensitivity, Lasso regression was slightly better, but the difference was not significant according to the adjusted  $\alpha$  [difference =  $-0.23, t(49) = -2.54, p = .014$ ]. As shown in Table 3, AUC reached .85 for random forest, which is generally considered as good predictive performance (45).

Figure 1 shows the importance measures from the two ML methods (i.e., random forest, Lasso regression) for the main evaluation center. For illustration purposes, we highlight the top five important predictors from each method. Recall that we trained random forest and Lasso regression to maximize each of the four criteria. Thus, there were four sets of importance scores for both methods. These importance scores for each method were highly consistent. Therefore, we reported the averaged importance scores across the four criteria for simplicity.

**Variable Importance.**—The values of VIMs from random forest were not correlated with those from Lasso regression (Person  $r = .15, p = .44$ ). However, their rank orders were positively related (Kendall's tau =  $.36, p = .02$ ). As mentioned above, we focus on the predictors with top five importance scores for our illustration. As expected, there is both overlap and discrepancy in terms of the top list in each of the methods. Three of the top five predictors were consistent across the methods, including past work years, biweekly work hours, and age. In addition, random forest identified total number of training hours and change in hourly wage as one of the influential predictors, while Lasso regression included position title (outreach/community positions vs. residential services) in the top five list.

The marginal effects of the top five predictors are visualized by PDPs in Figure 2. The general trends show that more past work years, older in age, or longer average work hours were associated with a lower probability of turnover. In contrast, a greater total number of training hours was associated with a higher probability of turnover. The random forest model also suggested that more change in hourly wage over the year was associated with a higher probability of turnover, while Lasso regression indicated that working in the outreach/community department vs. residential services was associated with a higher probability of turnover. In addition, as shown in Figure 2, the patterns from random forest appeared to be more nonlinear or nonmonotonic in some cases. For example, regarding the effect of past work years, random forest showed that the turnover probability steeply dropped within four years of hire while slightly yet gradually increased afterward.

### Results for the Application Center and the ML Methods' Applicability

The results of the application center, including the mean AUC, sensitivity, specificity, and overall accuracy, as well as their standard deviations from the 50 sets of predictions via the repeated K-fold cross-validation, are shown in Table 3. Paired t-tests suggested that random forest performed substantially better than Lasso regression for all criteria except for specificity: AUC (difference = 0.06,  $t(49) = 10.01, p < .01$ ), sensitivity (difference = 0.14,  $t(49) = 15.19, p < .01$ ), and overall accuracy (difference = 0.13,  $t(49) = 16.89, p < .01$ ). Lasso regression led to a higher specificity (difference = 0.14,  $t(49) = 15.19, p < .01$ ) than random forest, which means that Lasso regression performed better in terms of the true

negative rate (i.e., classifying stayers). The results of AUC above .8 indicate that the ML methods had a good prediction power for the application center's data.

Figure 3 presents the VIMs from both methods for the application center. Following the presentation of the main evaluation center result, we highlight the top five predictors. Both methods identified past work years as the most important predictor. Other than past work years, Lasso regression did not select any other variables, likely due to its penalty term in the discrepancy function (i.e., shrink effects/coefficients to zero among those with less impact) (51). Random forest also included hourly wage, age, biweekly work hours, and marital status in its top five list.

Three of the predictors (i.e., past work years, work hours, age) selected by random forest were consistent across both centers. However, the PDPs for these predictors suggest that the relationship patterns for some corresponding predictors were not necessarily consistent between the two centers (see Figures 2 and 4). For instance, increase in work hours were generally associated with lower probability of turnover in the main evaluation center, while it could increase the probability of turnover (e.g., after passing 70 hours biweekly) in the application center. Further, the turnover probability steeply dropped within the fourth year of hire (similar to the main evaluation center), but the probability sharply increased first before leveling off for the application center. A more dynamic pattern of age influence also seemed to be present for the application center. In addition to the three predictors, the results from the application center also showed that increase in wage may not always decrease the probability of turnover (e.g., it could increase the probability among those who are at the lowest wage range group) and married employees were more likely to leave.

### **Feasibility Assessment of HR Data Extraction for ML**

The brief survey returned by the HR director indicated high feasibility. Based on the director's report, on average, about one or two HR personnel were involved in each data extraction. Depending on the data management systems, the initial data extraction took about an hour to several days. The lengthiest process was determining the data fields that needed to be extracted between HR personnel and the research team (e.g., understanding available data fields and deciding on data format). For instance, the research team requested to extract employees' annual salaries; however, the annual salary varied depending on the data extraction point (e.g., varying work hours), and many employees were non-exempt. It required several days to write a program for the HR data management system to estimate an hourly rate for each employee at different time points (while managing their regular HR job duties). The levels of ease/difficulty of the data extractions were rated as either neutral or easy by the director once the data extraction fields were determined.

### **Discussion**

The goal of the current study was to examine the feasibility and predictive capacity of machine learning (ML) methods for community mental health center employees' turnover using historical HR data. We applied widely used parametric and nonparametric ML approaches with cross-validations to improve the generalizability of our findings. Our developed ML predictive models achieved a good level of predictive accuracy using the HR

data, particularly with random forest (the nonparametric approach) – both AUC and overall prediction accuracy were above or close to .8, considered good in general (45). In addition to the overall predictability, our ML methods were able to explore some key variables that may be important to predict turnover.

Because our primary purpose was to apply ML methods to estimate an individual employee's turnover probability given their available HR data (i.e., data-driven approach), not to determine predictors that are generalizable at the wider population level (i.e., theory-driven and hypothesis-testing approach), our interpretations of the predictors are limited or restricted to the specific organizations under the study. Regardless, for our illustration purpose, we highlighted the top five predictors that may be more influential (among the predictors considered) to turnover probabilities.

The top five list included past work years, wage, work hours, age, job position, training hours, and marital status at the two centers. These factors have been identified as important in general turnover research (52, 53). Our ML applications to HR data also mirrored the findings in the literature that the effects of turnover predictors or their directionality (i.e., positive vs. negative association with turnover) vary across studies. Such discrepancies in the literature may be due to differences in populations, study conditions, and predictors in their modeling. The two centers in our ML applications identified differences in predictors or their directionality, although some common predictors were also found (i.e., past work years, wage, age, work hours). While the two centers were similar in size and the mental health services they provided, one center was located in an urban area, and the other was in a rural area with different demographic decompositions in their workforce. The list of available predictors for ML was also different across the two centers. For instance, while the wage was highlighted as relatively important for both centers, the findings are not directly comparable, given the different conditions. We extracted all the historical wage data over multiple years from the main evaluation center, accounting for the potential salary inflation over the years. On the other hand, we only extracted the most recent wage data from the application center, which did not account for the salary inflation effect or change of the average wage over years. The current study focused on the initial feasibility testing of the ML approaches with limited implementation settings (e.g., the number of centers considered, predictors used, and sample size). To understand the difference in important predictors for turnover prediction across different centers, further applications of the current ML approaches at multiple locations, with differing contexts are needed as the next step.

Between the two ML approaches, random forest had better predictability compared to Lasso regression in terms of AUC and overall accuracy by a large margin, although Lasso regression could lead to higher specificity or sensitivity depending on the data. As mentioned in Wolper and Macready (54), no algorithm may outperform another in all metrics. In addition, the variable importance scores from the two approaches did not agree, although their rank orders were correlated. These differences are likely because some predictors might have nonlinear relationships with turnover, which were not captured by Lasso regression. Previous research showed the potential nonlinear dynamic nature of turnover decision-making processes (17). The two methods also differed regarding how the predictor redundancy was handled (55, 56). Random forest may be more robust to predictor

redundancy than Lasso regression because it has the feature to randomly select a subset of predictors in the tree building process. Because redundant predictors may not always be selected together, random forest can mitigate the impact of predictor redundancy. However, it is not immune to the problem. For example, Kubus (56) found that presence of redundant predictors could decrease the classification accuracy of random forest. Random forest is also less affected by outliers, given that it does not have any distributional assumptions. The bottom line is that the important predictors identified by either approach might warrant future investigation until one is certain about their prediction power and influential mechanisms. The influential predictors identified across multiple approaches would likely be more generalizable than those uniquely identified by a certain approach.

Some limitations of the study are worth mentioning. First, the current sample size was considered relatively small in the ML literature (57). Our intent for the current study was to pilot the feasibility of ML approaches for predicting turnover with historical HR data at community mental health centers. We chose decently sized centers in two states, with different geographic settings (urban and rural) that provided similar types of services. However, some of the categorical data were still sparse (although we decreased the sparseness by combining categories), and some of the continuous data had a substantial amount of missing data. We acknowledge that some community mental health centers may have fewer resources than our sites. For example, some of the HR data might be recorded in hard copy and some small centers may not have enough personnel to record all data electronically. In addition, it may be challenging for some small centers to customize data fields for ML if they do not have their own internal data management systems (e.g., payroll software), relying on external contractors. To resolve the issue, we will need to further train the ML predictive models by extending the number of application and testing centers. By doing so, we should be able to identify the core data and predictor format that may be important in general turnover prediction and those that may be important for specific organizational settings. Sequentially building ML models may be one approach to achieve this goal (e.g., build models with more generalizable predictors first, then add organization-specific predictors to improve the prediction). We will need to further investigate the strategies to borrow accumulated learning information when applying the ML approaches to smaller centers.

Second, as a preliminary study, we only examined two representative ML methods for classification problems. There are many more available ML methods to consider. For instance, for parametric methods, there are elastic net (as an extension of Lasso regression) and neural network. For nonparametric methods, gradient boosting machine (another tree-based nonparametric method) and support vector machine could be used (44). More complicated ML algorithms (e.g., deep learning) also could be used to fully explore the capacity of HR and administration data. As a fast-advancing area, new ML methods are continuing to be developed to overcome limitations of existing methods. The current feasibility study opens the possibility for using more powerful ML methods when a bigger sample is available.

Third, despite the good predictive performance of our ML predictive models, we recognize that HR data alone may not fully address the mechanisms of turnover. For instance, Fukui et

al. (58) found that some employee characteristics, such as demographics (that are typically available in historical HR data), could be predictive of actual turnover. Yet, other job factors, including job stressors (e.g., burnout, job satisfaction, perception of work-life conflict) may be more predictive for increased turnover *intention* among community mental health employees. Given the increased attention of mental health organizations on their employees' mental health, many organizations conduct organization-wide job well-being surveys to address job concerns. Such organizational data also could be integrated into the ML models, which may further improve the prediction accuracy.

Fourth, interpretations of important predictors or the directionality of their effects on turnover can be challenging, especially with random forest, when applying them to the actual employee management strategies without additional investigations. Unlike the traditional statistical approaches, which test pre-specified or theory-based hypotheses (specific variable relationships) in the population, the process of nonparametric ML (including random forest) often remains in a black box (59, 60). For instance, one predictor's effect (the relative strengths and directionality) may change depending on other predictor combinations. Such nonlinear relationships may be graphically explored with random forest (48, 61), yet the interpretation is not straightforward or easy with many predictors in ML. Additionally, output interpretations are restricted to the quality of data (e.g., garbage in, garbage out), or the data considered in the modeling. For example, our model indicated that a greater total number of training hours was associated with a higher turnover probability. Meaningful job training can be an important job advancement opportunity, facilitating employee retention. Therefore, the quality of training may matter (62), which was not considered in our models. The amount of required training hours also may be different across different job roles and interact with other factors. Therefore, further explorations are needed for HR and leadership to use the information for their turnover prevention and retention strategies. For instance, evaluating the characteristics of high turnover risk employees who might be clustered in a specific job, personal, or relational cluster may be informative for intervention purposes (although this would require sufficient data for the specific cluster).

Finally, but most importantly, despite the emerging popularity of ML approaches, the application to non-research data, including HR and administration data, requires caution. If the training data are systematically biased, the models trained based on the data will reproduce structural biases in our prediction models (25). For example, if specific racial groups systematically received biased job evaluations or were assigned to limited job roles that may be linked to turnover, ML predictive models could be biased. Data validity can also be a concern for data collected for non-research purposes (63). Synergy between data-driven approaches and theory-based approaches is suggested (64–66), which will be the important next step of the current study.

Despite the limitations, the current study provides new perspectives and avenues to address the persistent turnover issue among community mental health centers. Although still primitive, our study provides a data-driven approach to apply ML methods to existing HR and administration data, which may help address employee turnover. As the ML applications are accumulated across organizations, it may be expected that some

findings might be more generalizable across different organizations while others may be more organization-specific (idiographic). The former findings may contribute to broader policy and workforce development efforts. On the other hand, the latter findings can be useful for the organization's HR and leadership to evaluate and address turnover in their specific organizational contexts. As our continuing effort, it is important to study how the ML methods and outputs can be meaningfully utilized in routine management and leadership practice settings in mental health (including how to develop organization-tailored intervention strategies to support and retain employees) beyond identifying high turnover risk individuals. Such organization-based intervention strategies with ML applications can be accumulated and shared by organizations, which will facilitate the evidence-based learning communities to address turnover. This, in turn, may enhance the quality of care we can offer to clients.

To conclude, we would like to bring attention to ethical considerations of the proposed approach. Our ML methods are intended to identify high risk employees for turnover, given their HR and administration data, by calculating the individual turnover probability. This is analogous to the approaches in the modern medicine (67, 68), where we hope to identify high risk individuals for the organization to support them in preventing their potential future turnover. This is one of the essential HR efforts in data collection and use (e.g., addressing equity in gender, race, age). However, the method could be used against high turnover risk employees (e.g., similar to the current insurance systems excluding or burdening individuals with a high health risk probability). It is important to continue considering how our ML approaches can help mental health employees, but not add further barriers for successful job retention.

## Acknowledgements

The study was supported by National Institute of Mental Health R34 (NIMH R34MH119411).

## References

1. Aarons GA, Sawitzky AC. Organizational climate partially mediates the effect of culture on work attitudes and staff turnover in mental health services. *Administration and Policy in Mental Health and Mental Health Services Research* 2006;33(3):289–301. [PubMed: 16544205]
2. Beidas RS, Marcus S, Wolk CB, Powell B, Aarons GA, Evans AC, et al. A prospective examination of clinician and supervisor turnover within the context of implementation of evidence-based practices in a publicly-funded mental health system. *Administration and Policy in Mental Health and Mental Health Services Research* 2016;43(5):640–9. [PubMed: 26179469]
3. Bukach AM, Ejaz FK, Dawson N, Gitter RJ. Turnover among community mental health workers in Ohio. *Administration and Policy in Mental Health and Mental Health Services Research* 2017;44(1):115–22. [PubMed: 26662638]
4. Hyde PS. Report to Congress on the Nation's substance abuse and mental health workforce issues. Rockville, MD: Substance Abuse and Mental Health Services Administration. 2013.
5. Thomas KC, Ellis AR, Konrad TR, Holzer CE, Morrissey JP. County-level estimates of mental health professional shortage in the United States. *Psychiatric Services* 2009;60(10):1323–8. [PubMed: 19797371]
6. Hawkins M 2021 Review of Physician and Advanced Practitioner Recruiting Incentives. White Paper, available at <https://www.merrithawkins.com/uploadedFiles/physician-advanced-practitioner-incentive-review-2021.pdf>. 2021.

7. Eby LT, Burk H, Maher CP. How serious of a problem is staff turnover in substance abuse treatment? A longitudinal study of actual turnover. *Journal of Substance Abuse Treatment* 2010;39(3):264–71. [PubMed: 20675097]
8. Knudsen HK, Ducharme LJ, Roman PM. Clinical supervision, emotional exhaustion, and turnover intention: A study of substance abuse treatment counselors in the Clinical Trials Network of the National Institute on Drug Abuse. *Journal of Substance Abuse Treatment* 2008;35(4):387–95. [PubMed: 18424048]
9. Hoge M, Morris J, Daniels A, Stuart G, Huey L, Adams N. An action plan for behavioral health workforce development Cincinnati, OH: Annapolis Coalition on the Behavioral Health Workforce. 2007.
10. Holtom BC, Burch TC. A model of turnover-based disruption in customer services. *Human Resource Management Review* 2016;26(1):25–36.
11. Yanchus NJ, Periard D, Moore SC, Carle AC, Osatuke K. Predictors of job satisfaction and turnover intention in VHA mental health employees: A comparison between psychiatrists, psychologists, social workers, and mental health nurses. *Human Service Organizations: Management, Leadership & Governance* 2015;39(3):219–44.
12. Rollins AL, Salyers MP, Tsai J, Lydick JM. Staff turnover in statewide implementation of ACT: Relationship with ACT fidelity and other team characteristics. *Administration and Policy in Mental Health and Mental Health Services Research* 2010;37(5):417–26. [PubMed: 20012481]
13. Aarons GA, Sommerfeld DH, Hecht DB, Silovsky JF, Chaffin MJ. The impact of evidence-based practice implementation and fidelity monitoring on staff turnover: evidence for a protective effect. *Journal of Consulting and Clinical Psychology* 2009;77(2):270–280. [PubMed: 19309186]
14. Cho YJ, Song HJ. Determinants of turnover intention of social workers: Effects of emotional labor and organizational trust. *Public Personnel Management* 2017;46(1):41–65.
15. Scanlan JN, Still M. Job satisfaction, burnout and turnover intention in occupational therapists working in mental health. *Australian Occupational Therapy Journal* 2013;60(5):310–8. [PubMed: 24089982]
16. Dąderman AM, Basinska BA. Job Demands, engagement, and turnover intentions in polish nurses: the role of work-family interface. *Frontiers in Psychology* 2016;7:1621. [PubMed: 27847481]
17. Holtom BC, Mitchell TR, Lee TW, Eberly MB. 50 years of turnover and retention research: a glance at the past, a closer review of the present, and a venture into the future. *Academy of Management Annals* 2008;2(1):231–74.
18. Rombaut E, Guerry M-A. Predicting voluntary turnover through human resources database analysis. *Management Research Review* 2018;41(1):96–112.
19. Raza A, Munir K, Almutairi M, Younas F, Fareed MMS. Predicting Employee Attrition Using Machine Learning Approaches. *Applied Sciences* 2022;12(13):6424.
20. Kelleher JD, Mac Namee B, D'arcy A. *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies*: MIT press; 2020.
21. Mishra SN, Lama DR, Pal Y. Human Resource Predictive Analytics (HRPA) for HR Management in Organizations. *International Journal of Scientific & Technology Research* 2016;5(5):33–5.
22. Sikaroudi E, Mohammad A, Ghousi R, Sikaroudi A. A data mining approach to employee turnover prediction (case study: Arak automotive parts manufacturing). *Journal of Industrial and Systems Engineering* 2015;8(4):106–21.
23. Xie Y, Li X, Ngai E, Ying W. Customer churn prediction using improved balanced random forests. *Expert Systems with Applications* 2009;36(3):5445–9.
24. Ribes E, Touahri K, Perthame B. Employee turnover prediction and retention policies design: a case study. *arXiv preprint arXiv:170701377* 2017.
25. Sajjadiani S, Sojourner AJ, Kammeyer-Mueller JD, Mykerezi E. Using machine learning to translate applicant work history into predictors of performance and turnover. *Journal of Applied Psychology* 2019;104(10):1207–1225. [PubMed: 30907603]
26. Quinn A, Rycraft JR, Schoech D. Building a model to predict caseworker and supervisor turnover using a neural network and logistic regression. *Journal of Technology in Human Services* 2002;19(4):65–85.

27. Kuncel NR, Klieger DM, Connelly BS, Ones DS. Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis. *Journal of Applied Psychology* 2013;98(6):1060–72. [PubMed: 24041118]
28. Kuhn M, Johnson K. *Feature engineering and selection: A practical approach for predictive models*: CRC Press; 2019.
29. Kuhn M, Johnson K. *Applied predictive modeling*: Springer; 2013.
30. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, et al. Missing value estimation methods for DNA microarrays. *Bioinformatics* 2001;17(6):520–5. [PubMed: 11395428]
31. van Mens K, Kwakernaak S, Janssen R, Cahn W, Lokkerbol J, Tiemens B. Predicting future service use in Dutch mental healthcare: A machine learning approach. *Administration and Policy in Mental Health and Mental Health Services Research* 2022;49(1):116–24. [PubMed: 34463857]
32. Couronné R, Probst P, Boulesteix A-L. Random forest versus logistic regression: a large-scale benchmark experiment. *BMC bioinformatics* 2018;19(1):1–14. [PubMed: 29291722]
33. Pereira JM, Basto M, Da Silva AF. The logistic lasso and ridge regression in predicting corporate failure. *Procedia Economics and Finance* 2016;39:634–41.
34. Gareth J, Daniela W, Trevor H, Robert T. *An introduction to statistical learning: with applications in R*: Springer; 2021.
35. Hastie T, Friedman JH, Tibshirani R. *The elements of statistical learning: data mining, inference, and prediction*: Springer; 2001.
36. Muchlinski D, Siroky D, He J, Kocher M. Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data. *Political Analysis* 2016;24(1):87–103.
37. Liu M, Wang M, Wang J, Li D. Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B: Chemical* 2013;177:970–80.
38. Breiman L, Friedman J, Olshen R, Stone C. *Classification and regression trees*. Wadsworth & Brooks. *Cole Statistics/Probability Series* 1984.
39. Kirasich K, Smith T, Sadler B. Random forest vs logistic regression: binary classification for heterogeneous datasets. *SMU Data Science Review* 2018;1(3):9.
40. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* 2009;14(4):323–348. [PubMed: 19968396]
41. Breiman L. Random forests. *Machine learning* 2001;45(1):5–32.
42. Kuhn M. Building predictive models in R using the caret package. *Journal of Statistical Software* 2008;28:1–26. [PubMed: 27774042]
43. Hastie T, Tibshirani R, Friedman J. *Unsupervised learning*. *The elements of statistical learning*: Springer; 2009. p. 485–585.
44. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning: With Applications in R*: Springer; 2021.
45. Hosmer DW Jr, Lemeshow S, Sturdivant RX. *Applied logistic regression*: John Wiley & Sons; 2013.
46. Kuhn M, editor. *Variable Importance Using The caret Package* 2007.
47. Loh W-Y, Zhou P. Variable importance scores. *arXiv preprint arXiv:210207765* 2021.
48. Molnar C. *Interpretable machine learning*: [Lulu.com](https://lulu.com); 2020.
49. Friedman JH. Greedy function approximation: a gradient boosting machine. *Annals of Statistics* 2001:1189–232.
50. Team RC. *R: A language and environment for statistical computing* R Foundation for Statistical Computing, Vienna, Austria. 2012. 2021.
51. Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 1996;58(1):267–88.
52. Cotton JL, Tuttle JM. Employee turnover: A meta-analysis and review with implications for research. *Academy of Management Review* 1986;11(1):55–70.

53. Tsai SP, Bernacki EJ, Lucas LJ. A longitudinal method of evaluating employee turnover. *Journal of Business and Psychology* 1989;3(4):465–73.
54. Wolpert DH, Macready WG. No free lunch theorems for optimization. *IEEE transactions on evolutionary computation* 1997;1(1):67–82.
55. Archer KJ, Kimes RV. Empirical characterization of random forest variable importance measures. *Computational Statistics & Data Analysis* 2008;52(4):2249–60.
56. Kubus M The problem of redundant variables in random forests. *Acta Universitatis Lodzianis Folia Oeconomica* 2018;6(339):7–16.
57. van der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. *BMC Medical Research Methodology* 2014;14(1):1–13. [PubMed: 24383436]
58. Fukui S, Rollins AL, Salyers MP. Characteristics and job stressors associated with turnover and turnover intention among community mental health providers. *Psychiatric Services* 2020;71(3):289–92. [PubMed: 31658895]
59. Gonzalez MF, Capman JF, Oswald FL, Theys ER, Tomczak DL. “Where’s the IO?” Artificial intelligence and machine learning in talent management systems. *Personnel Assessment and Decisions* 2019;5(3):5.
60. Leavitt K, Schabram K, Hariharan P, Barnes CM. Ghost in the machine: On organizational theory in the age of machine learning. *Academy of Management Review* 2021;46(4):750–77.
61. Cafri G, Bailey BA. Understanding variable effects from black box prediction: Quantifying effects in tree ensembles using partial dependence. *Journal of Data Science* 2016;14(1):67–95.
62. Menne HL, Ejaz FK, Noelker LS, Jones JA. Direct care workers’ recommendations for training and continuing education. *Gerontology & Geriatrics Education* 2007;28(2):91–108. [PubMed: 18032193]
63. Xu H, Zhang N, Zhou L. Validity concerns in research using organic data. *Journal of Management* 2020;46(7):1257–74.
64. Maass W, Parsons J, Purao S, Storey VC, Woo C. Data-driven meets theory-driven research in the era of big data: Opportunities and challenges for information systems research. *Journal of the Association for Information Systems* 2018;19(12):1.
65. Hoffer JG, Ofner AB, Rohrhofer FM, Lovri M, Kern R, Lindstaedt S, et al. Theory-inspired machine learning—Towards a synergy between knowledge and data. *Welding in the World* 2022:1–14.
66. Nelson LK. Computational grounded theory: A methodological framework. *Sociological Methods & Research* 2020;49(1):3–42.
67. Quiroz-Juárez MA, Torres-Gómez A, Hoyo-Ulloa I, León-Montiel RdJ, U’Ren AB. Identification of high-risk COVID-19 patients using machine learning. *Plos One* 2021;16(9):e0257234. [PubMed: 34543294]
68. Bottrighi A, Pennisi M, Roveta A, Massarino C, Cassinari A, Betti M, et al. A machine learning approach for predicting high risk hospitalized patients with COVID-19 SARS-Cov-2. *BMC Medical Informatics and Decision Making* 2022;22(1):1–13. [PubMed: 34983500]

**Implications for Health Care Provision and Use:**

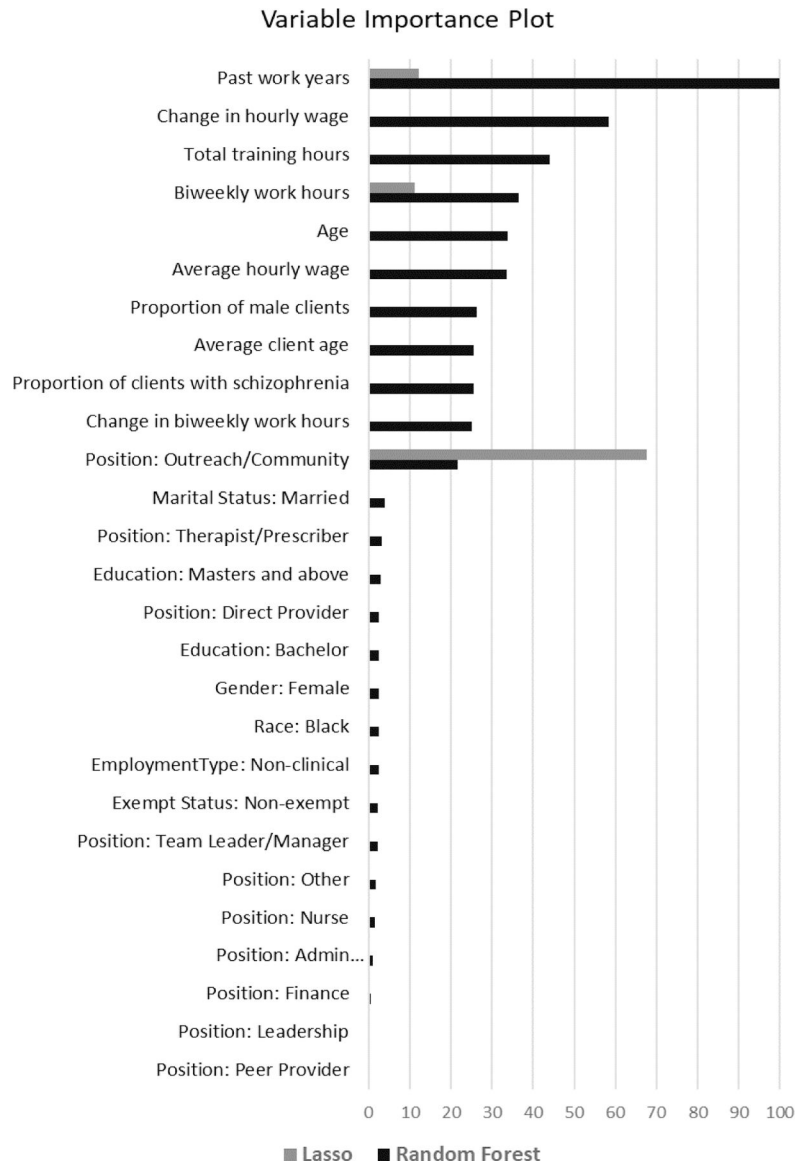
The organization-specific findings can be useful for the organization's HR and leadership to evaluate and address turnover in their specific organizational contexts. Preventing extensive turnover has been a significant priority for many mental health organizations to maintain the quality of services for clients.

**Implications for Health Policies:**

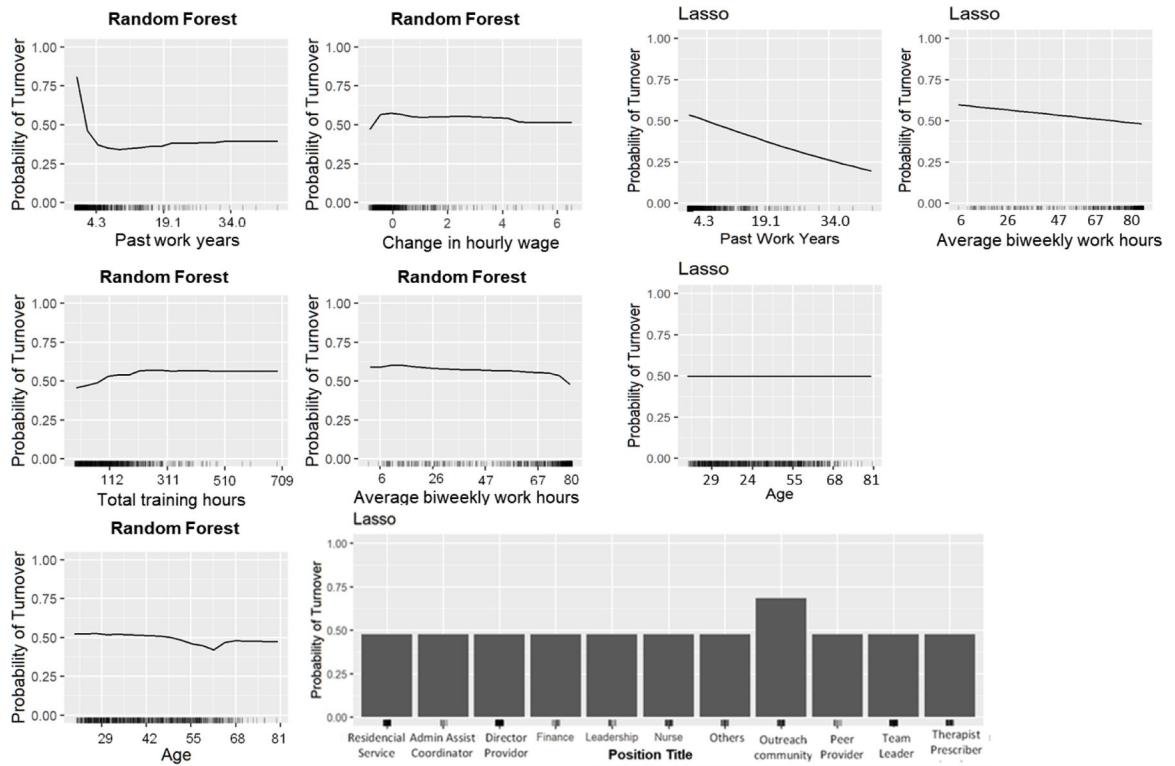
The generalizable findings may contribute to broader policy and workforce development efforts.

**Implications for Further Research:**

As our continuing research effort, it is important to study how the ML methods and outputs can be meaningfully utilized in routine management and leadership practice settings in mental health (including how to develop organization-tailored intervention strategies to support and retain employees) beyond identifying high turnover risk individuals. Such organization-based intervention strategies with ML applications can be accumulated and shared by organizations, which will facilitate the evidence-based learning communities to address turnover. This, in turn, may enhance the quality of care we can offer to clients. The continuing efforts will provide new insights and avenues to address a data-driven, evidence-based turnover prediction and prevention strategies using HR data that are often under-utilized.

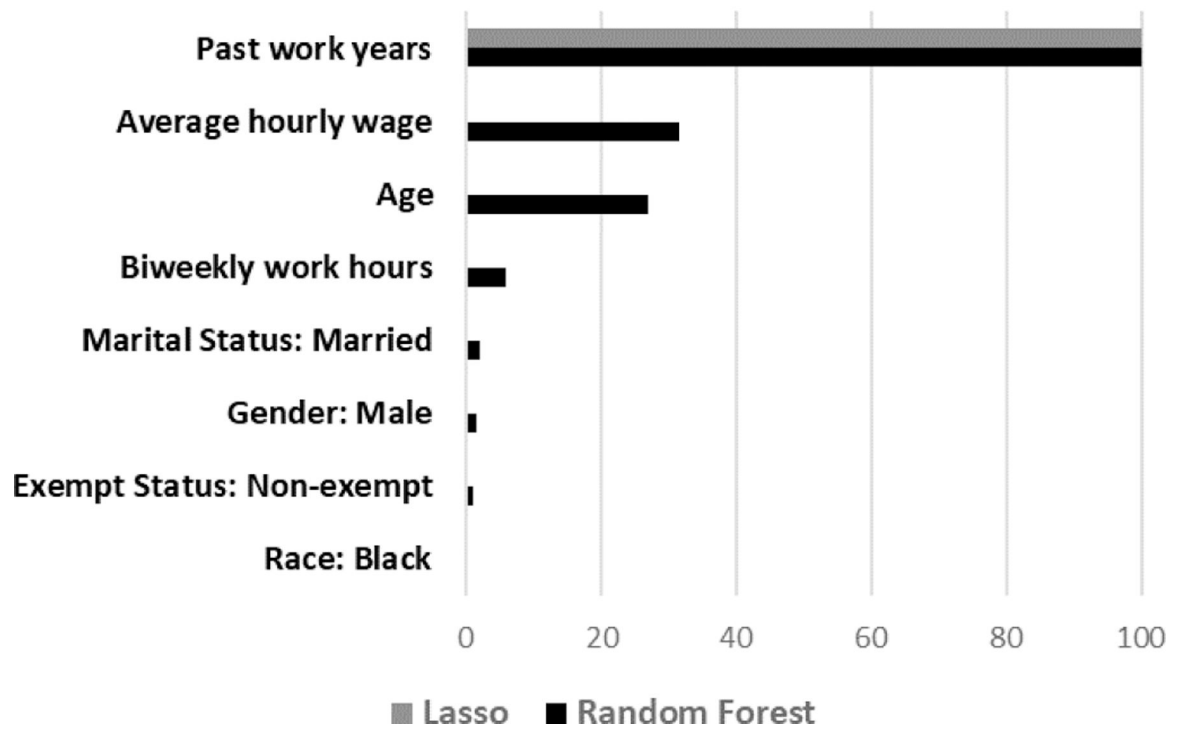


**Figure 1.** Variable Importance Plot from Lasso Regression and Random Forest for the Main Evaluation Center  
*Note:* MISS categories are not included. Because the Lasso function (to reduce effects/coefficients down to zero for predictors with less impact), the importance score of age (relative to other more influential predictors) is very small (.05), thus not apparent within the graph; however, it was the last predictor of the top five list from Lasso regression.



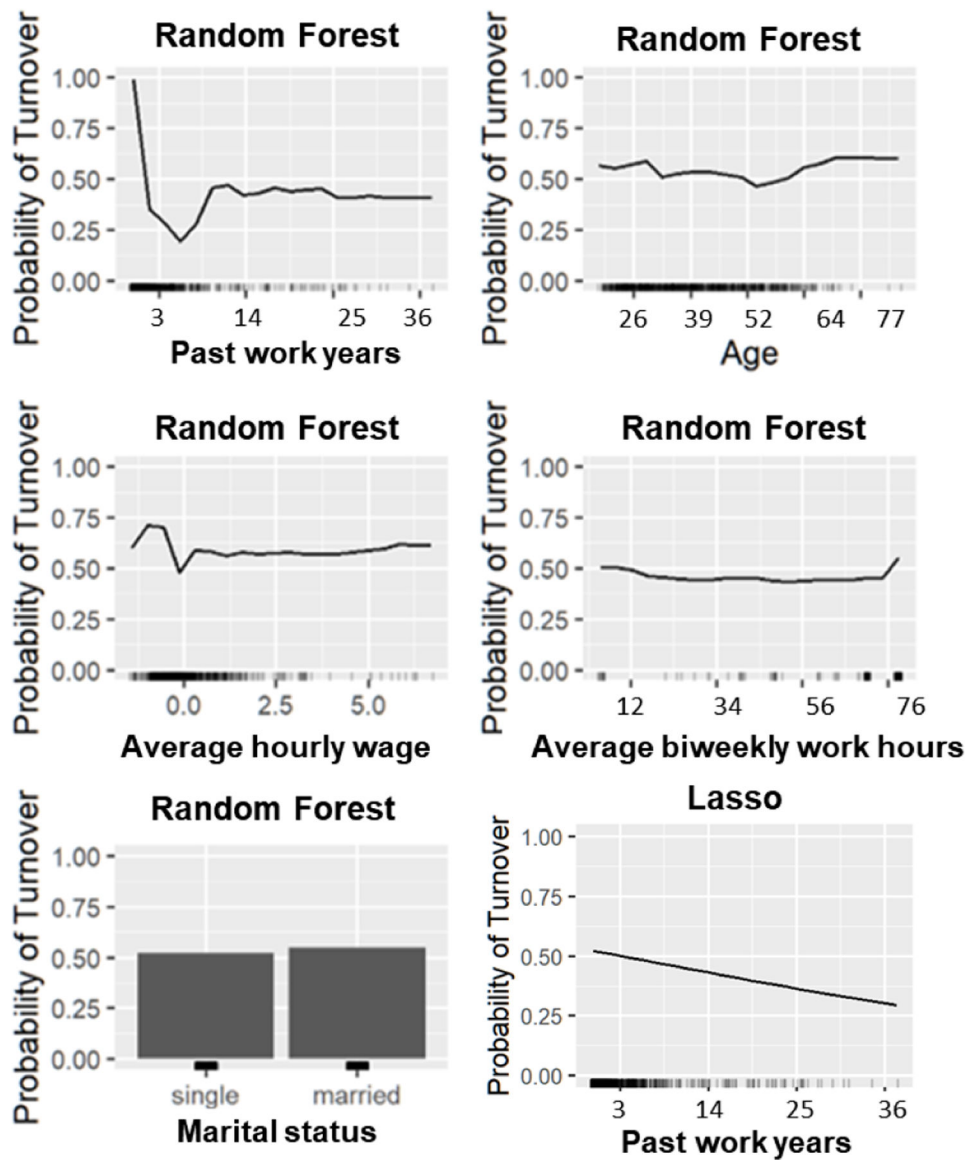
**Figure 2.** Partial dependence plots (PDPs) of the top five predictors from each method for the main evaluation center  
*Note.* The PDPs portray marginal effects of the predictors. The y axis represents the predicted probability of turnover based on the fitted model. Average hourly wage is in a standardized log transformed unit.

### Variable Importance Plot



**Figure 3.** Variable Importance Scores from Lasso Regression and Random Forest for the Application Center

*Note:* MISS categories are not included.



**Figure 4.** Partial dependence plots (PDPs) of the top five predictors for the application center  
*Note.* Average hourly wage is in a standardized log transformed unit.

**Table 1**

Frequency Distribution Of Categorical Predictors For Stayers And Leavers From The Two Centers.

	<u>Main Evaluation Center</u>		<u>Application Center</u>	
	<b>Stayer (n = 318)</b>	<b>Leaver (n = 336)</b>	<b>Stayer (n = 407)</b>	<b>Leaver (n = 487)</b>
<b><u>Marital Status</u></b>				
Married	109	93	181	213
<i>Single</i>	<i>204</i>	<i>235</i>	<i>162</i>	<i>198</i>
Missing	4	8	64	76
<b><u>Education</u></b>				
<i>Associate and below</i>	<i>101</i>	<i>95</i>		
Bachelor	93	122		
Masters and above	124	119		
<b><u>Exempt Status</u></b>				
<i>Exempt</i>	<i>62</i>	<i>52</i>	<i>81</i>	<i>47</i>
Non-exempt	256	284	326	440
<b><u>Race</u></b>				
Black	138	147	9	19
<i>White</i>	<i>177</i>	<i>172</i>	<i>360</i>	<i>413</i>
Asian/Pacific islander	3	0	20	30
Missing	0	17	14	15
<b><u>Employment Type</u></b>				
<i>Clinical</i>	<i>249</i>	<i>251</i>		
Non-clinical	69	47		
Missing	0	38		
<b><u>Gender</u></b>				
Female	237	247	327	396
<i>Male</i>	<i>76</i>	<i>88</i>	<i>80</i>	<i>91</i>
Missing	5	1	0	0
<b><u>Position</u></b>				
Admin Assist/Assistant/coordinator	10	4		
Direct Provider	106	78		
Finance	14	6		
Leadership	11	10		
Nurse	16	22		
Outreach/Community	6	45		
Peer Provider	9	8		
<i>Residential Service</i>	<i>39</i>	<i>62</i>		
Team Leader/Manager	60	26		
Therapist/Prescriber	28	21		
Other	19	12		
Missing	0	42		

Note. The reference category for each categorical predictor is in italics.

**Table 2**

The Descriptive Statistics of Continuous Predictors For Stayers And Leavers From Two Centers

Predictors	Main Evaluation Center Stayer (n = 318)					Main Evaluation Center Leaver (n = 336)						
	n	mean	sd	median	min	max	n	mean	sd	median	min	max
Age	318	44.2	13.1	43.2	22.1	77.1	336	40.4	12.6	38.0	21.4	80.5
Past work years	318	5.6	6.2	3.4	0.1	44.4	336	3.2	5.5	1.1	0.0	39.5
Hourly wage <sup>a</sup>	317	0.3	0.1	0.2	0.0	1.1	333	0.3	0.1	0.2	0.0	1.1
Average biweekly work hours	316	71.5	17.4	79.4	7.8	80.0	332	63.7	22.6	76.7	2.3	80.0
Average client age	192	46.7	11.7	48.8	9.5	83.0	164	48.1	9.0	49.2	13.1	63.1
Proportion of male clients	192	0.6	0.2	0.6	0.0	1.0	164	0.6	0.2	0.6	0.0	1.0
Proportion of clients with schizophrenia	192	0.5	0.3	0.5	0.0	1.0	164	0.5	0.3	0.5	0.0	1.0
Change (standard deviation) of hourly wage <sup>b</sup>	316	0.0	0.1	0.0	0.0	0.3	272	.04	.04	0.02	0.0	0.3
Change (standard deviation) of biweekly work hours	314	4.5	7.4	1.2	0.0	35.1	272	6.3	8.5	2.3	0.0	40.0
Total training hours	299	120.6	112.5	93.0	1.0	683.9	310	103.1	84.6	85.9	0.8	443.7

Predictors	Application Center Stayer (n = 407)					Application Center Leaver (n = 487)						
	n	mean	sd	median	min	max	n	mean	sd	median	min	max
Age	407	41.2	13.0	39.3	19.7	85.8	487	36.6	12.4	34.9	17.9	83.8
Past work years	407	5.0	6.2	2.4	0.6	37.2	487	1.8	4.1	0.3	0.0	29.3
Hourly wage <sup>c</sup>	407	0.2	0.2	0.2	0.0	1.1	487	0.2	0.1	0.2	0.0	1.0
Average biweekly work hours	407	76.1	12.1	80.0	4.0	80.0	487	78.3	9.8	80.0	4.0	80.0

Note.

<sup>a</sup>Hourly wage was log transformed due to its non-normality and standardized by year

<sup>b</sup>Standard deviation (SD) of log transformed hourly wage

<sup>c</sup>Hourly wage was log transformed.

**Table 3**

The Overall Predictive Performance of The Machine Learning Approaches

Criterion	Main Evaluation Center		Application Center	
	Lasso Regression <sup>I</sup>	Random Forest <sup>I</sup>	Lasso Regression <sup>I</sup>	Random Forest <sup>I</sup>
<b>AUC</b>	.76(.04)	.85(.03)	.83(.03)	.89(.03)
<b>Sensitivity</b>	.87(.05)	.85(.05)	.45(.06)	.95(.02)
<b>Specificity</b>	.47(.06)	.73(.05)	.89(.03)	.75(.05)
<b>Accuracy</b>	.67(.04)	.78(.03)	.70(.04)	.82(.03)

*Note.* The values outside the parentheses are means and inside the parentheses are standard deviations. AUC = area under the curve.

<sup>I</sup>The means and standard deviations of the criteria for random forest and Lasso regression are obtained based on the corresponding hyperparameter value tuned to maximize each criterion.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript