Distribution Agreement

Signature:

Kimberly N. Gajewski 4/25/2013

Demographic Trends of Sick Leave Absenteeism among Civil Service Employees at the Centers for Disease Control and Prevention from 2004-2012

By

Kimberly N. Gajewski

Master of Public Health

Epidemiology

Kevin Sullivan, PhD

Committee Chair

Ross Spears, PhD

Committee Member

Demographic Trends of Sick Leave Absenteeism among Civil Service Employees at the Centers for Disease Control and Prevention from 2004-2012

By

Kimberly N. Gajewski

B.A., B.S.

Duke University

2011

Thesis Committee Chair: Kevin Sullivan, PhD

An abstract of

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2013

Abstract

Demographic Trends of Sick Leave Absenteeism among Civil Service Employees at the Centers for Disease Control and Prevention from 2004-2012

By Kimberly N. Gajewski

For organizations such as the Centers for Disease Control and Prevention (CDC), understanding trends and variations in sick leave use within the workforce is essential for the development of surge capacity plans and targeted leave messaging, which may ultimately contribute to a higher attendance rate and healthier workforce. We analyzed sick leave use among CDC civil servants working in Atlanta, GA between 2004-20012 by demographic variables including age, gender, length of service, and pay grade, as well as CDC-specific variables including emergency response tier qualification and retirement plan. Then we used a mixed methods approach and Type III analysis to build a descriptive model of sick leave proportions and demographic variables.

Sick absenteeism usage varied significantly (variation of greater than 1 sick day per year) by gender, EOC response tier, length of service at the CDC, age, and GS pay grade level. Women took on average 2 full days of sick leave per year more than men. Participants between 35-44 years old took the most sick leave of any age group. Those among the higher response tiers took significantly less sick leave than those among the lower response tiers. Further, the proportion of sick leave taken by those among the highest response tier was significantly lower than the population's average proportion of sick leave (5.3 days per year vs. 7.3 days per year).

Our final descriptive model contained age, gender, response tier and an interaction term between age and gender. While younger women tended to have lower proportions of sick leave than men in the same age category, women between 45-54 years old had significantly higher proportions of sick leave than men in the same age category controlling for age and response tier qualification.

This study was the first of its kind to examine the relationship between demographics and absenteeism at the CDC, and provides an initial stepping stone for further investigation into these complex associations. Future studies should examine these associations on smaller time scales, perhaps breaking the data down by month or even day of the week.

Demographic Trends of Sick Leave Absenteeism among Civil Service Employees at the Centers for Disease Control and Prevention from 2004-2012

By

Kimberly N. Gajewski

B.A., B.S.

Duke University

2011

Thesis Committee Chair: Kevin Sullivan, PhD

An abstract of

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2013

Acknowledgements

I would first like to thank Captain Ross Spears and Dr. Carrie McNeil for introducing this project to me and sticking with me through an endless series of obstacles and edits. This project would have been impossible without their unending support and guidance. I would also like to thank Dr. Paul Meechan for helping me think about the "why" behind my results.

Next I would like to thank Remi Oyinloye for assembling my dataset and working with me to get the data presented exactly as it needed to be for my analysis. I would also like to thank my faculty advisor, Dr. Kevin Sullivan for helping me create my sick leave proportion variable and for remaining calm when I was panicking about SAS.

Finally I would like to thank my parents for always believing in me and supporting me through two years of graduate school. I promise to get a job now.

Table of Contents

I. Background/Literature Review

Evaluating and understanding trends in sick and total absenteeism allows organizations to develop strategic policies, such as continuity of operations and capacity loss mitigation, during instances of higher absenteeism. In order to develop these policies, companies and organizations need to collect and routinely analyze data on absenteeism among the total workforce as well as by demographics and other characteristics.

Founded in 1946, the United States Centers for Disease Control and Prevention (CDC) is tasked with being able to launch a rapid response to public health disasters or emergencies such as foodborne outbreaks, tornadoes or pandemics. The main CDC campus is located in Atlanta, GA but there are 12 other locations scattered throughout the United States and territories. The CDC employs approximately 15,000 people, including 8,000 contractors and 1,000 U.S. Public Health Service Commissioned Corps (PHS) officers.

This study evaluates temporal absence trends among civil service employees at the CDC between 2004-2012 accounting for specific demographics such as age, gender, pay grade, and length of service. Predicting sick leave absenteeism trends among specific workforce subpopulations could provide managers with a mechanism to pre-identify times when staffing levels may not be adequate to address agency mission requirements. In addition, results of this analysis could help target sick leave messaging to specific populations within the workforce.

The literature assessing longitudinal trends in total or sick leave absenteeism is diverse in the world of business. However, very little of it pertains to longitudinal trends in federal government agencies. Historically, most of the literature on employee absenteeism has been conducted in private or corporate businesses, as these types of organizations were initially more interested and capable of funding research oriented towards maximizing the output and retention of their workforces (Farrell 1984; Leonard,Vanameringen et al. 1987; Rantanen and Tuominen 2011). Recently, there has been an increasing interest in absenteeism as it relates to general workforce statistics from both private and federal agencies within the United States and across the globe. Budget tightening, in addition to maximizing worker output, is one impetus for the rising interest. Several agencies are in the process of collecting and reporting on employee attendance trends, though there have been few publications (van Vuuren, de Jong et al. 2008).

Numerous approaches involving a variety of variables have been employed to assess sick leave and total absenteeism. Generally, these variables can be divided into three broadly descriptive categories; individual, group and organizational predictors. The first category of variables involves individual predictors, which are demographic descriptors. These predictors have the longest history of use in peer-reviewed literature because they are fairly easy to obtain and analyze (Leonard, Vanameringen et al. 1987; Vandenheuvel and Wooden 1995). Individual predictors include gender, age, race, family size, education level, marital status, health status, and mental health status. Many researchers have explored how these variables play a role in determining why a person goes to work or remains at work (Farrell 1984; Leonard,Vanameringen et al. 1987; Vandenheuvel and Wooden 1995).

The second category of variables involves group predictors, which are variables that assess worker interactions within the company. Group predictors include assessments of the quality of interpersonal office relationships, office sizes and layout, and rates of absenteeism amongst cohorts of coworkers (Leonard, Vanameringen et al. 1987; Hilton, Sheridan et al. 2009). The third category of variables involves organizational predictors, which are sometimes difficult to explicitly define because these predictors can broadly impact entire companies, regions, or even countries. Historically, organizational predictor variables include economic health analysis, company attendance policies, retirement plans, and sick leave policies (Wilson and Peel 1991; Farrell 1984; Hendrix and Spencer 1989; van Vuuren, de Jong et al. 2008).

Variable selection is heavily dependent on the scope of the study, availability (or obtainability) of employee data, and overarching interests of the investigator. The following sections elaborate on each of the categories of variables and provide examples of conclusions that have been drawn from them.

Individual Predictors

Individual predictors of employee absenteeism have been utilized in a large number of studies across the globe with varying conclusions. While some individual predictors are inexpensive and easy to obtain, others require considerably greater investments to collect.

For example, variables that assess physical health, the development and/or persistence of chronic illnesses, mental well-being, and personal job satisfaction levels on a longitudinal basis may be difficult to obtain due to privacy laws and often expensive to acquire. In comparison, demographic predictors such as gender, age and length of service are easier and less expensive to obtain. Unsurprisingly, these demographic variables are the most often assessed in absenteeism studies. However, despite the fact that these data are relatively ease to obtain, they can be difficult to interpret, and can lead researchers to widely different or even opposing conclusions. These opposing conclusions may be a result of variations in organizational policies, such as offering paid sick leave, and can make generalizing study findings to other populations very difficult.

Gender

Age and gender are two of the most widely studied demographic variables in absenteeism literature (Leonard, Dolan et al. 1990; Harrison and Martocchio 1998). The association between level of absenteeism and gender has been studied for decades (Leigh 1983). Some studies, especially those conducted when women first began to enter the workforce in significant numbers, concluded that women had significantly higher levels of absenteeism than men (Leigh 1983; Meisenheimer 1990). These researchers also found that women between the ages of 25-34 had the highest frequency of absences and concluded that this was due to the fact that women in this age bracket were much more likely to have small children that needed care (Meisenheimer 1990). The presence of preschool children in the household made women twice as likely to have an absence from work than women who had older children or no children at all (Meisenheimer 1990).

However, the presence of dependent children in the household did not influence absence rates among either gender in the study conducted by Vandenheuvel and Wooden (1995).

Some researchers have not found gender to be a relevant predictor of sick or total absenteeism. For example, a UK study by Sharp and Watt (1995) found that women were more likely to take absences of one day, while men were more likely to take absences between 2-6 days long. Even though women took more frequent absences, their absences were shorter in length—thus, both genders were absent for approximately the same total amount of time (Sharp and Watt 1995). The earlier studies that concluded women had greater levels of absence evaluated frequency, but not total length of absence.

AGE

As with the studies of gender, the research on the association between age and absenteeism is inconclusive. Meisenheimer (1990) found that workers between 25-54 years old generally had the lowest rates of absenteeism when not stratified by gender. Sharp and Watt (1995) found that although the number of absences tends to decrease with the increase of age, the length of absences tends to increase with age for both older men and women.

Findings on the interaction between age, gender and absenteeism have arrived at contrasting conclusions. A meta-analysis conducted by Martocchio (1989) concluded the frequency of absence steadily declines with employee age. In addition, several studies have found that the youngest and oldest age groups were the most likely to be absent

(Leonard, Vanameringen et al. 1987; Hendrix and Spencer 1989; Martocchio 1989; Leonard, Dolan et al. 1990; Vandenheuvel and Wooden 1995). While Sharp and Watt (1995) found no difference in absenteeism between men and women, Martocchio (1989) found a stronger trend of declining absenteeism among men than women. He concluded that age and absenteeism might be completely unrelated among women (Martocchio, 1989).

The relationship between age and absenteeism may also be influenced by numerous other variables besides gender. For example, younger professionals may have lower job satisfaction and, therefore, may be absent more frequently (Martocchio 1989). Older professionals are more likely to have higher-paying and highly specialized jobs, which may make them more reluctant to miss work (Hendrix and Spencer 1989). Older professionals, more likely to have worked for the company for a longer period of time, may also have higher company loyalty (Harrison and Martocchio 1998). However, younger employees may take fewer absences to “prove” themselves to the company or get a strong start in their career (Sharp and Watt 1995).

LENGTH OF SERVICE

Length of service has been found to be negatively associated with absenteeism (Bernadin 1977), not associated with absenteesim (Nicholson, Brown and Chadwick-Jones 1977), and positively associated with absenteeism (Baumgarten and Sobol 1979) in various studies. These conflicting conclusions may suggest that this variable is particularly prone to the influence of other variables in the working environment. For example, employees that are happier at their workplace may be less likely to take absences and may be more likely to stay at that job for a greater length of time.

Nearly half of all turnover occurs during the first year of an employee's time with the organization (Farrell 1984). Older employees with longer tenure may exhibit lower rates of absenteeism because of a better person-organization fit that emerges over time (Harrison and Martocchio 1998). Oshagbemi (2000) found that job satisfaction increases progressively with length of service, which may contribute to lower absenteeism. Job satisfaction also increases with job continuity; therefore, employees that have worked for the same agency over an extended period of time may have fewer absences than those who have moved from position to position in other agencies (Oshagbemi 2000).

As with previous risk factors, there may be several effect modification variables between length of service and absenteeism. One noted confounding factor is if the absences were paid or if they were unpaid (Farrell and Peterson 1984). Gender is another potential modifier. Although absenteeism is generally found to be highest amongst the newest workers, one study found that female employees who were employed for less than a year were found to have lower absenteeism rates than other women (VandenHeuvel and Wooden 1995).

Pay Grade

The General Schedule (GS) scale is used to pay all full-time civil service United States government employees (OPM, 2012). While the GS rates for each city are adjusted according to living costs, this standardized system allows for reasonably accurate conclusions about job types across the agency. Employees at lower grades, such as those between 3-7, have clerical-type jobs, while those at grades above 8 usually have more managerial roles(Hendrix and Spencer 1989). GS 13 employees and above represent management and executives (Hendrix and Spencer 1989). Typically higher pay grades reflect higher positions within the agency and, often, greater responsibility and seniority. Hendrix and Spencer (1989) developed a multivariate model of absenteeism among employees working at the U.S. Department of Defense. They found pay grade was significantly associated with absenteeism. Employees with higher pay grades were less likely to be absent than those who had lower pay grades (Hendrix and Spencer 1989). Similarly, Lu et al. (2010) found that socioeconomic status, when considered with employment conditions, had a larger impact on predicting sickness-based absence than gender. Those of a lower socioeconomic status (SES) were much more likely to have sick-related absences and to have longer sick-related absences than those of a higher SES (Lu, Laaksonen et al. 2010). This study, however, did not evaluate differences in health care availability or overall state of health between the high and low SES groups.

Interestingly, Hendrix and Spencer (1989) also found pay grade reduced the effects of burnout and cold/flu episodes on total absenteeism. Employees with higher pay grades were less likely to be absent, even if they were experiencing burnout or were ill. The authors propose this may be due to their perceived level of importance. People working in healthcare services have been shown to have an increased likelihood of going to work while sick (Rantanen and Tuominen 2011). The most common barriers to taking sick leave among this demographic include: difficulty in replacement, amount of make-up work that must be covered after the absence, and attitudes about their own level of health. However, attendance at work while being ill could have negative impacts on productivity and may ultimately lead to greater rates of absenteeism in that office, if other employees contract the illness.

Group Predictors

Researchers commonly incorporate group-level predictors into their absenteeism studies. These predictors include assessments of the quality of interpersonal office relationships, office sizes and layout, and rates of absenteeism amongst cohorts of coworkers. One benefit of using these types of variables is that they allow researchers to examine absence as a cultural phenomenon within work groups. These variables assess the customs and practices of how, when, and why work groups are present or absent at given times.

Within most office environments, taking leave can depend heavily on the quality of interpersonal office relationships and the perceptions of coworkers (Mason and Griffin 2003.) However, Goldberg and Waldman (2000) investigated the role job satisfaction plays in employee absenteeism rates within a hospital in the northeastern United States. Despite previous studies that concluded job satisfaction was related to absenteeism, Goldberg and Waldman (2000) found no association. In addition, job satisfaction did not play an intermediary role between individual predictors, such as marital status, number of children, or position level, and level of absenteeism. One study of New Zealand public employees discovered that the rate of absenteeism between different groups over time depended on the group’s size (Dew, Keefe and Small 2005). They concluded that larger groups had higher rates of absenteeism than smaller groups because there was less accountability and sense of community in the larger groups (Dew, Keefe and Small 2005).

Organizational Predictors

Organizational predictors of absenteeism have not been as extensively studied as individual and group predictors. These variables are difficult to draw conclusions from, because their specific impact on every individual is often unknown. That being said, several studies have identified organizational predictors that were significantly associated with absenteeism.

At the organizational level, researchers have found that during times of economic downturn absenteeism rates go down (Wilson and Peel 1991). The 2005 Bureau of Labor Statistics reported an absenteeism rate of 2.9% among healthcare workers or those working in healthcare related industries in the private sector. The 2010 U.S. Census Bureau reported an annual average sickness absence rate of 2.2% overall (1.8% for men, 2.8% women). The unfavorable economic environment may have played a substantial part in the reduction in absenteeism. However, declining rates of absenteeism were not as common among employees in the federal sector, possibly reflecting the fact that civil service positions are more stable and offer benefits that are less commonly provided in the private sector.

Other organizational-level predictors of absenteeism trends include company policies that reward attendance or offer paid leave (Wilson and Peel 1991). Overall, having paid sick leave increases absence, though absence rates vary due to a variety of factors independent from the sick leave policy (Harrison and Martocchio 1998; Rantanen and Tuominen 2011). Although some research has been published on the organizational and societal benefits from implementing paid leave policies (Farrell 1984; Hendrix and Spencer 1989; van Vuuren, de Jong et al. 2008), studies have found that paid sick leave reduces the incidence of illness in the workplace and actually diminishes total absenteeism in the office (van Vuuren, de Jong et al. 2008).

Retirement plan

Civil servants working for the CDC may be under one of two different retirement systems, the Civil Servant Retirement System (CSRS) or the Federal Employee Retirement System (FERS). Federal workers under the Federal Employees Retirement System (FERS) are able to count their unused sick leave as creditable service in calculating their retirement annuity, whereas under CSRS it cannot be counted. One possible outcome, which has not yet been studied, is that employees working under the FERS system may feel incentivized to come in to work while sick instead of taking sick leave. Aronsson, Gustafsson and Dallner (2000) found that one-third of employees surveyed in welfare and education-related fields in Sweden reported coming in to work despite being ill. In general, people whose occupations require them to provide care or education to others, or people whose occupations rely on interpersonal office relationships were more likely to come to work sick (Aronsson 2000). Perceptions of low “replaceability” also contribute to sick presenteeism. Those who reported that they would have to make-up virtually all of the work that they missed were less likely to take sick days (Aronsson 2000).

Another group of researchers found that those with jobs that provided services to people were found to be less likely to take sick leave because they felt a responsibility to their clients (Caverley, Cunningham & MacGregor 2007). In addition to this sense of responsibility, the investigators found that employees in a Canadian public service organization were less likely to take sick days when there was an increased sense of job insecurity and increased workplace demands (Caverley, Cunningham & MacGregor 2007). Several studies, such as those by Aronsson et al. (2000), Caverly et al. (2007), and Dew et al. (2004), have illustrated the negative impact sick presenteeism can have on the productivity and overall functioning of a workforce.

In one study population, consistently going to work while ill contributed to an increased incidence of serious coronary events (Dew, Keefe, Small 2005). Aronsson et al. (2000) observed that organizations with high sick presenteeism ultimately had higher rates of absenteeism because ill employees were spreading their illness to other employees.

CDC-Specific Predictor

For organizations such as the CDC, predicting absences allows for the ability to build alternate surge capacity in their staffing plans. Surge capacity planning may involve planning for external contracting, reevaluating leave usage particularly among leadership, or drafting alternative internal response plans that do not rely on the absent staff to be present.

Emergency Operations Center responder tier

The bioterrorism mission statement of the CDC includes the phrase, "…to lead the public health effort in enhancing readiness to detect and respond to bioterrorism." The Emergency Operations Center (EOC) is tasked with ongoing 24/7 surveillance of public health situations throughout the country and around the world. It serves as the central point of contact for all US health agencies to report potential threats, as well as the main operations center for coordinating a response to a public health emergency anywhere in the world.

Every employee within the CDC can register to work in the EOC during a response. In fact, all CDC staff members are highly encouraged to complete the online profile detailing their skills and qualifications for responding to public health emergencies. Based upon the skills they list, employees are assigned a response tier level. EOC staff members utilize tier levels as one way to quickly assign the most appropriate employees to a response.

Registering with the EOC is entirely optional, and some civil servants choose not to sign up. For tracking purposes, these people are considered to be part of Tier 0, which means they will not be called upon to assist with an emergency response. Civil servants who are part of Tier 1 have registered with the EOC and completed introductory training;

however they are highly unlikely to be called in to assist with a response. Tier 2 responders have completed more response training than people in Tier 1. They are called to fill critical positions within the EOC during a response operation. Responders in Tier 3 are capable of being deployed to the field during a response. Finally, those assigned to Tier 4 are capable of filling response leadership or liaison positions both within the EOC and out in the field (see Table 1 for further elaboration). Employees who qualify as Tier 4 responders may also be dually registered as Tier 2 or Tier 3 responders, if they are willing to work on a response in any capacity necessary. Any predictive trends in absenteeism based upon response tier qualification could have significant implications for recruiting and managing incident responders within the CDC.

Table 1. Emergency Response Tier Levels and Responsibilities

CDC Responder Tiers	Duties during a response
Tier 0	Not registered
Tier 1	Registered, unlikely to assist in a response
Tier 2	Fill critical positions in the EOC during a response
Tier 3	Deploy to the field during a response
Tier 4	Hold a leadership or liaison role either in the EOC or out in the field during a response

Previous work has been done examining sick leave (Spears et al. 2013) and total leave (paper in progress) absenteeism within this population, but neither paper examines how leave trends differ by individual and group predictors. The researchers used 10 years of data to establish a baseline sick leave level in order to build a predictive model that specified trigger points of sick leave levels higher than expected by the baseline. Specifically, a trigger event was defined as an exceedance in the value of percent sick leave for a time point above the sum of the mean and a fixed number of standard deviations for that time point. An evaluation of total leave provided a mechanism to identify seasonal trends in absences so that the CDC as a whole could predict times of lower staffing capacity.

One of the study's most interesting results was that approximately 14.5% of staff were not available to work at any given moment. That equates to roughly 1 in 7 employees. This absenteeism rate is nearly double that of the general population, as assessed by the Current Population Survey. During the holidays, the absenteeism rate was found to be even higher- nearly 35%.

These projects were able to predict to a fine degree the percentage of sick and total leave rates at any given time of year at the organizational level. The researchers found that attendance dips predictably during flu season, in late September/early October shortly after school starts, during the spring, and towards the end of the year when some people may use their leftover sick days as vacation time. In order to better understand these predictable trends in absenteeism, this study will take the important next step of understanding how demographics and pay grade may influence those leave trends.

This study is the first at the CDC to determine if there are any associations among Tier level qualification and sickness absenteeism. If there are periods of time during which staff from leadership ranks are all absent, the agency may be less able to rapidly and appropriately respond to public health emergencies during those times.

Predicting trends in sick absenteeism levels allows organizations to develop strategic staffing and internal response policies that ensure their agency is capable of functioning and fulfilling their mission mandates during all situations. These findings could prompt targeted sick leave use messaging, or absence policy and procedural changes that ensure the agency mandates can be accomplished even with the loss of personnel.

III. Manuscript

Demographic Trends of Sick Leave Absenteeism among Civil Service Employees at a Federal Public Health Agency from 2004-2012

Gajewski, Kimberly BS,BA

Abstract

For organizations such as the Centers for Disease Control and Prevention (CDC), understanding trends and variations in sick leave use within the workforce is essential for the development of surge capacity plans and targeted leave messaging, which may ultimately contribute to a higher attendance rate and healthier workforce. We analyzed sick leave use among CDC civil servants working in Atlanta, GA between 2004-20012 by demographic variables including age, gender, length of service, and pay grade, as well as CDC-specific variables including emergency response tier qualification and retirement plan. Then we used a mixed methods approach and Type III analysis to build a descriptive model of sick leave proportions and demographic variables.

Sick absenteeism usage varied significantly (variation of greater than 1 sick day per year) by gender, EOC response tier, length of service at the CDC, age, and GS pay grade level. Women took on average 2 full days of sick leave per year more than men. Participants between 35-44 years old took the most sick leave of any age group. Those among the higher response tiers took significantly less sick leave than those among the lower response tiers. Further, the proportion of sick leave taken by those among the highest

response tier was significantly lower than the population's average proportion of sick leave (5.3 days per year vs. 7.3 days per year).

Our final descriptive model contained age, gender, response tier and an interaction term between age and gender. While younger women tended to have lower proportions of sick leave than men in the same age category, women between 45-54 years old had significantly higher proportions of sick leave than men in the same age category controlling for age and response tier qualification.

This study was the first of its kind to examine the relationship between demographics and absenteeism at the CDC, and provides an initial stepping stone for further investigation into these complex associations. Future studies should examine these associations on smaller time scales, perhaps breaking the data down by month or even day of the week.

Introduction

Evaluating and understanding trends in sick absenteeism use within an organization allows that organization to develop strategic policies, such as continuity of operations and capacity loss mitigation, during instances of higher absenteeism. In order to develop these policies, companies and organizations need to amass and assess key sets of data on worker absenteeism, as well as demographic and other characteristic information about the workforce.

The literature assessing longitudinal trends in total or sick leave absenteeism is diverse; however, very little of it pertains to longitudinal trends in federal agencies. Historically, most of the literature on employee absenteeism has been conducted in private or corporate businesses, as these types of organizations were initially more interested and capable of funding research oriented towards maximizing the output and retention of their workforces (Farrell 1984; Leonard, Vanameringen et al. 1987; Rantanen and Tuominen 2011). However, recently there has been an increasing interest in absenteeism as it relates to general workforce statistics from federal agencies within the United States and across the globe.

There are numerous approaches, and numerous variables, that have been employed to assess sick leave and total absenteeism. Generally however these variables can be divided into three broadly descriptive categories. The first category of variables involves individual predictors, which are the true demographic descriptors of a person. These predictors have the longest history of use in peer-reviewed literature because they are fairly easy to obtain and analyze (Leonard, Vanameringen et al. 1987; Vandenheuvel and Wooden 1995). Individual predictors include gender, age, race, family size, education level, marital status, health status, and mental health status. Many researchers have explored how these variables play a role in determining why a person goes to work or remains at work (Farrell 1984; Leonard,Vanameringen et al. 1987; Vandenheuvel and Wooden 1995).

The second category of variables involves group predictors, which are variables that assess worker interactions within the company. Group predictors include assessments of the quality of interpersonal office relationships, office sizes and layout, and rates of absenteeism amongst cohorts of coworkers (Leonard, Vanameringen et al. 1987; Hilton, Sheridan et al. 2009). The third category of variables involves organizational predictors, which are difficult to explicitly define because these predictors broadly impact entire companies, regions, or even countries. Historically studied organizational predictor variables include economic health analysis, company attendance policies, retirement plans, and sick leave policies (Wilson and Peel 1991; Farrell 1984; Hendrix and Spencer 1989; van Vuuren, de Jong et al. 2008).

Variable selection is heavily dependent on the scope of the study, availability of employee data, and overarching interests of the investigator. The following sections elaborate on each of the categories of variables and provide examples of conclusions that have been drawn from them.

This study examines the longitudinal trends of absenteeism amongst CDC full-time civil service employees by using individual and organizational-level demographic variables to build a predictive mathematical model of sick and total absenteeism levels. Additionally, the study will break down absenteeism trends by each variable individually; to identify predictable patterns of leave among specified components of the workforce.

All data, which includes employee information from 2004-2012, have been harvested from an automated online attendance tracking system (TASNet). Data are collected only from civil service full-time employees within the agency because other employees, such as those working for the Public Health Service, report their attendance in a separate system. However, the agency's 7,000+ civil servants comprise over 50% of the 14,000-person workforce. We have made the assumption they are representative of the workforce as a whole.

The data were considered exclusively in aggregate to ensure privacy. All demographic information was de-identified. Although the agency maintains several facilities of varying sizes, only those agency locations within the Atlanta, Georgia metropolitan area were used for this analysis due to the large sample size provided at this location.

Variables of interest were identified either based on existing literature or the importance of a given variable to organizational preparedness. The variables that will be explored in this study include: total leave, sick leave, age, gender, retirement plan (Civil Service Retirement System or Federal Employees Retirement System), General Schedule (GS) pay level, length of service, and Emergency Operations Center (EOC) Tier qualification. The use of demographic data is critical due to the evidence of their association with absenteeism in existing literature.

Analyzing the association between EOC tier qualification and/or GS level and absenteeism could lead to the identification of times of decreased mission capability as well as limited response capacity. Since the CDC is charged with being able to mount a timely response to any public health emergency, the discovery of any decreased internal response capacity could be a major issue. Understanding the relationship between these variables and sick leave allows for improved personnel planning in order to address specific population groups, particularly during periods of high illness such as peak flu season. Understanding the temporal trends of total leave rate of key employee groups will allow better preparedness and response planning by understanding when these groups are and are not normally available during the year. This study will describe specific and consistent trends of absenteeism levels over time by some or all of the above listed variables, and will provide a predictive model of these trends.

Null Hypotheses

1. Sick absenteeism proportions will not vary by age, gender, EOC response Tier, length of service at the CDC and GS pay grade level.
2. Employees enrolled in the Federal Employee Retirement System (FERS) plan will not have significantly lower sick leave proportions than those enrolled in the Civil Service Retirement System (CSRS).

3. A predictive model incorporating employee GS level, length of service, and EOC Tier qualification cannot be developed to predict timing of sick leave absenteeism among critical employees.

Methods

The study population consists of all full-time civil service employees working at the CDC Atlanta campus between 2004-2012. Other CDC employees, such as fellows, interns and Public Health Service officers were not included because their attendance is tracked through other systems. Full time civil service employees at other CDC campus locations outside of Atlanta were excluded because those campuses are significantly smaller, which increases the likelihood of being able to identify individuals in the dataset. Employees at other campuses were also excluded because there are noticeable seasonal variations in sick leave depending on geographic location and that would have caused a significant amount of noise during analysis.

This secondary dataset was assembled from time records reported in the CDC Time and Attendance System Network (TASNet). Employee time is recorded biweekly, and amount of time worked is reported in 15 minute increments. Although time is self-reported, it must be approved by the individual's supervisor; therefore, it is considered highly accurate though the system does allow for retroactive adjustments. This study will be examining trends in sick leave. Sick leave is accrued at a constant rate of 4 hours per pay period, and can be used by employees for illness or for dental and medical appointments for themselves or family members.

Staff from the CDC Management Information Systems Office (MISO) extracted time information from the TASNet system for all Atlanta civil service employees from 2004-2012. This time information included total amount of time the employee could have possibly worked, the actual amount of time the employee worked, the amount of sick leave taken, and the amount of total leave (which includes all types of paid and unpaid leave) taken. All of these variables were given in hours.

MISO staff linked demographic information from Human Resources and time records for each of individual. Demographic information included the variables gender, age, pay grade, length of service, response tier qualification, and retirement system type. All of the records were randomly assigned ID numbers, and no names were included in any step to further ensure that the individuals are unidentifiable.

Prior to being granted access to this secondary dataset from MISO, we applied for and were granted an exemption from the CDC IRB (Appendix A). Emory University's IRB has a standing agreement with the CDC IRB that they will honor the CDC's decision, and this project was also exempted by the Emory IRB.

The variables age, pay grade, and length of service were entered categorically instead of continuously to provide additional assurance that information would not be individually identifiable. Age was reported in nine five-year categories, beginning with 20-24 years old and ending with 60 years and older. Pay grade (as defined using the General Schedule scale) was reported in four categories, 01-03, 04-07, 08-11 and 12-15. Length of service was reported in seven categories, the first two ranges (0-3 years and 4-9 years) were selected specifically to capture the changes in leave allowances that occur after five years and ten years with the CDC. The subsequent five categories were standard five-year increments starting with 10-14 years and ending with 30 or more years. Full descriptive statistics for these variables are available in Appendix B.

Initial data descriptive analysis was conducted in SAS (10.0) to identify incomplete records. Any records that were missing information including gender, pay grade, length of service, responder tier qualification, retirement system or age were deleted from the dataset. Prior to permanent deletion, the removed records were compared to the retained records to check for biased deletions. The removed records reflected the data diversity observed in the retained records, therefore there was no deletion bias.

Once cleaning was completed, several of the variables were re-coded using Microsoft Excel and SAS (10.0). The response tier qualification variable contained some records that were numeric and some records that were text-based. For example, some records were listed as “2 and 3”, indicating that the individual was qualified for responding in both capacity levels. Due to the relatively small number of dually qualified individuals, their data were recoded to reflect qualification in only the higher of the two tiers. For example, someone qualified to respond in a tier 2 and 4 capacity was recoded to be qualified to respond in just a tier 4 capacity. The tier response qualification system is hierarchically structured; therefore, someone who is qualified as a tier 4 responder automatically is capable of responding at the tier 3 and 2 levels. Because of this structuring, it was unnecessary to retain the lower qualification tier in dually identified records.

Despite having the total number of hours each person worked and the total amount of sick leave they took within the study timeframe, we needed to create a variable that could account for the variation in amount of time each person worked during the study years. Not everyone worked all nine years, so assessing sick leave using a denominator of nine years would artificially bias results against those who worked the whole time. For example, if Person A and Person B both took thirteen sick days, but Person A worked two years and Person B worked eight years, then using a denominator of nine years they would appear equal, but using a proportional denominator, it becomes obvious that Person B has taken far less sick leave relative to their time worked. We devised a formula that would standardize sick leave use independent of actual time worked by turning it into a yearly proportion. This proportion is reflected by the following formula:

$$\text{Average number of sick days per year} = \frac{\text{Sick leave (days)}}{\text{Total possible days worked}} \times \text{Work days per year}$$

Basic statistics, such as cell frequency, mean and median sick leave values, as well as quartile values were calculated for all the variables using the frequency and univariate procedures in SAS (10.0). Despite the large size of this dataset, some of the categorical variables contained extremely small cells, which could impact analysis; therefore, we had to restructure them slightly. The ages were collapsed into four categories, 20-34, 35-44, 45-54 and 55+. The youngest age category is slightly larger than the others in order to roughly distribute the same number of people in each category. Some categories within the length of service variable were highly similar, so they were collapsed together, resulting in the following new categories 0-3, 4-9, 10-19, 20-29 and 30+. Similarly, the two highest pay grade categories were collapsed into one, encompassing grades 08-15.

Preliminary assessments of significance were conducted in SAS (10.0) using generalized linear models with Poisson distributions. Regular linear regression models were not applied to this data because it violates the normal distribution assumption. All predictor variables were assessed for significance individually using Type III contrast analysis within a generalized linear model using a significance level of $p < 0.0001$, which was chosen based upon our determination of meaningful significance which is discussed below.

Due to the large size of the dataset, we determined that traditional alpha significance values, such as 0.01 or 0.05 would not result in meaningful differences in our dataset. Therefore we created an additional test for significance based upon what we consider to be a meaningful difference in sick leave from a managerial perspective. An experienced manager within the CDC decided that a difference of more than one sick day per year would be considered meaningfully different. Therefore, all variables not only had to be statistically significant at the $p < 0.001$ level, their parameter estimates also had to differ from each other by more than a day per year in order for the variable to be retained in the model.

After each variable was tested individually, we created a full model containing all six of the predictor variables (age, gender, length of service, responder tier qualification, pay grade and retirement plan). This full model was used to assess all 15 possible interaction terms for significance. Each interaction term was tested in the full model independently, that is, only one interaction term was included at a time. All of the interactions that were not significant at the $p < 0.0001$ level were discarded from analysis. After all of the interactions were tested individually, the significant ones were put back into the full model all at the same time. The most insignificant interaction term was deleted and the full model was run again. This process was repeated until all of the interaction terms in the model were significant at $p < 0.001$.

The remaining interaction terms were then assessed for meaningful differences, that is, the parameter estimates were calculated to determine if they varied from each other by more than one sick day per year. If any terms were not different by greater than one sick day per year they were dropped from the model one at a time, beginning with the variable that had the least difference. Following this process, the only remaining interaction terms in the model are those that are significant at $p < 0.0001$ whose levels vary by greater than one sick day per year.

Finally, the predictor variables themselves were examined for significance based upon our two inclusion criteria. Any predictor variable that was involved in a significant interaction term was automatically retained in the model. Using this process, it is possible that the final model could contain one or more insignificant predictor variables, however they have to be retained in order for the model to be hierarchically well formulated. Any predictor variables (not involved in any of the interaction terms) that failed either significance test were dropped using the same process described above.

Throughout the elimination process we were careful to check for any confounding. If a variable's removal resulted in dramatic parameter estimate changes we retained it in the model and noted that it was a confounding factor.

Results

Demographic

Approximately 9% of the observations in the dataset were removed during data cleaning, dropping the total number of observations from 7,479 to 6,781. The most common reason observations were deleted was because they were missing response Tier qualification (N=295). All of the deleted observations did not otherwise significantly differ from the retained observations; therefore their removal did not result in any biases.

The majority of the population was female (64.5%) and a substantial portion worked at the GS 08-15 level (92.2%). Only 6.25% were participating in the CSRS retirement plan and the majority was not signed up for a response Tier (72.5%, see Table 1). The average number of sick days taken per year across the entire population was 7.3 sick days/year (95% CI 7.2, 7.5). Women took significantly more sick days per year than men (8 days/year vs. 6.1 days/year, $p<0.0001$, see Figure 2.). Variation between the age groups, though still significant (p<0.0001) was not as pronounced; 20-34yr olds took the least amount of sick leave (6.5 days/year) and 35-44yr olds took the most amount of sick leave (7.7 days/year). Refer to Figure 1 for a complete distribution of sick leave proportions among the age groups.

There was no statistical difference in sick leave between members in either the CSRS or FERS retirement systems (p=0.2). This finding rejects our second hypothesis that those under the FERS system will take significantly less sick leave per year than those under the CSRS system. Amount of sick leave varied significantly (p<0.0001) by length of service at the CDC, response tier qualification and GS pay grade. Those with a GS 01-03 took an average of 4.3 sick days/year while those with a GS 04-07 took an average of 10.2 sick days/year. Interestingly, there was a linear association between response Tier qualification and sick leave, with the lowest Tier taking the most sick leave (7.5 days/year) and the highest Tier taking the least sick leave (5.3 days/year). Refer to Table 1 and Figure 3 for complete statistics.

These findings support our first hypothesis that absenteeism proportions will vary by EOC response tier, length of service at the CDC and GS pay grade level. In addition, these findings suggest that there are also significant differences in absenteeism proportions based upon age and gender.

MODELING

We used a mixed methods approach to model selection. Our initial model contained all six of the study predictor variables. Pay grade, response tier qualification, age, gender and length of service were all statistically significant in our Type III analysis at the 0.0001 significance level. Retirement plan was the only insignificant variable in the model ($p=0.02$). Despite its insignificance, retirement plan was retained in the model because we wanted to determine if there were any confounding issues.

As described in the methods section, interaction terms were added to the model individually to assess for initial significance. Out of 15 possible interaction terms, 6 were statistically significant ($p<0.0001$, refer to Appendix C). Prior to adding the significant interaction terms back into the model, we checked some of the variables to ensure that their interactions would have enough individuals in each level (or cell). If there are very small cells, for example with less than 10 people in them, it could compromise our statistical model tests. We checked retirement system against Tier qualification and length of service because we were concerned that some response Tiers would not have anyone on the CSRS plan in them. In addition, we checked the age variable against the length of service variable, and the pay grade variable. Both of the variables we checked against retirement system contained small cell values, as well as the age variable checked against the length of service variable. Due to their small cell sizes, these three interaction terms were discarded from further assessment in the model.

The remaining three significant interaction terms (pay grade*age, pay grade*length of service, and age*gender) were simultaneously added to the initial model and we proceeded to use the backwards elimination method to reduce it. The first insignificant interaction term in the full model was pay grade*length of service ($p=0.007$).

After dropping that term and re-running the model, all of the remaining interactions were statistically significant ($p<0.0001$) and there were no dramatic changes in the parameter estimates. However, we still checked the interaction terms to make sure that they were meaningfully different (difference of greater than 1 sick day/year). The pay grade*age interaction term was not meaningfully different (difference = 0.88 sick days/year). When we re-ran the model without that interaction, the sole remaining interaction term (age*gender) remained statistically and meaningfully different.

The next step in our backward elimination process was to determine if any of the original 6 predictor variables could be dropped. By default, since the age*gender interaction term was statistically and meaningfully different; both the age and gender predictor variables had to be retained in the model. Removing either variable would result in a non-hierarchically well-formed model.

Retirement system, which was initially insignificant, remained insignificant ($p = 0.026$); and, therefore, it was the first predictor variable to be dropped. Removal of that variable did not greatly change the parameter estimates. All of the remaining variables were statistically significant ($p<0.0001$), but pay grade was not meaningfully different (difference =0.67 sick days/year) so it was the next to be dropped. In the subsequent version of the model, length of service was not meaningfully different (difference = 0.78 sick days/year), and it was dropped. Throughout this process we were careful to watch for any changes in parameter estimates as a result of dropping a variable. There were no significant changes in those estimates; therefore, we concluded that none of the variables we removed were confounders.

Our final model is shown below-

SICK DAYS TAKEN PER YEAR = AGE + GENDER + RESPONSE TIER + AGE*GENDER

The final parameter estimates are shown in Table 2. The negative linear association between sick leave proportion and response tier qualification remained after controlling for age and gender. Employees in the two lowest response Tiers took sick leave on average 8.8 days each year, while employees in the highest response Tier took on average 7 sick days per year. There is an interesting interaction between age and gender in our final model. Women aged 20-34 and 35-44 took approximately a day less sick leave than men aged 20-34 and 35-44. However, the association reversed for women aged 45-54, who took on average one sick day more per year than men in the same age category.

Discussion

During data cleaning, approximately 9% of our dataset were removed because those observations contained missing values on one or more of the study variables of interest. Because these observations did not differ from the retained observations, we concluded that their removal did not skew the results of our analysis.

Sick absenteeism proportions were observed to vary significantly (variation of greater than 1 sick day per year) by gender, EOC response Tier, length of service at the CDC, and GS pay grade level. The variation between all of the age categories was less pronounced; however, it was also observed to be significant. These findings support our first hypothesis that sick leave proportions vary significantly by five of our six key variables of interest.

Sharp and Watt (1995) concluded that while women took more frequent absences, they were shorter in duration, therefore the men and women in their study had approximately the same total absence time. Although our study was not able to assess the length or frequency of absences, we examined the total amount of sick leave taken and converted that number into a proportion of sick leave per year. If our population was similar to the one studied by Sharp and Watt (1995) we expected to observe no difference in sick leave proportions between male and female employees. However, our study concluded that women take on average 2 full days of sick leave per year more than men. Our conclusion is in agreement with the conclusions drawn by Leigh (1983) and Meisenheimer (1990). However, Meisenheimer (1990) went on to observe that women between 25-34 years old had the highest level of absenteeism, and that was not observed in our study. We concluded that women between 45-54 years old had the highest proportion of sick absenteeism of any age. The reasons behind this group's elevated proportion of absenteeism should be explored further through qualitative interviews or surveys.

Previous studies of the association between age and absenteeism level have concluded that younger workers, particularly the youngest workers, tend to have higher absenteeism levels than older workers (Harrison & Martocchio 1998; Oshagbemi 2000). Surprisingly, the youngest age group in our study (20-24 years old) took almost a day less of sick leave per year than the oldest age group (55+ years). Participants between 35-44 years old took the most sick leave, at an average of one day per year more than the oldest age group. These findings are difficult to interpret under the current literature because traditional explanations for higher or lower levels of absenteeism (such as higher job satisfaction for older employees and childcare responsibilities for younger employees) do not apply as well. Further research should be conducted to determine the factors behind this specific age group's sick leave use proportions. One possibility is that this group has had to start caring for their aging parents.

Participants in this study who worked at the CDC between 10-29 years had the highest proportion of sick leave absence. These participants took an average of 2.2 more sick days per year than participants who worked at the CDC between 0-3 years, and an average of one sick day more per year than participants who worked at the CDC between 4-9 or 30+ years. Harrison and Martocchio (1998) hypothesized that employees with longer tenure may exhibit lower proportions of absenteeism because of a better person-organization fit that emerges over time; however this does not appear to be an applicable factor in this population. Younger workers and those that have only been working at the CDC for a short period of time might have lower levels of sick leave due to high job enthusiasm or fear of being fired, or they may be saving it up.

The differences in sick leave proportions between various pay grade categories were the most striking in this study. Those in a pay grade between 04-07 took nearly 6 more sick days per year than those in a pay grade between 01-03, and 3 more sick days per year than those in a pay grade between 08-15. Hendrix and Spencer (1989) observed that pay grade was significantly associated with absenteeism for Department of Defense employees. They found that employees with higher pay grades were less likely to be absent than those who had lower pay grades. However, this trend was not observed in our study because the lowest absenteeism proportions were observed in the lowest pay grades.

Interestingly, we observed a strong gradient effect in proportion of sick leave and response Tier qualification. Participants who were qualified under the highest response Tier (Tier 4) took an average of 5.3 sick days per year, which is well below the overall population average proportion of 7.3 sick days per year. Those qualified as Tier 1 responders, or not qualified for a response Tier at all (Tier 0), had the highest proportion of sick leave at 7.5 sick days per year. This gradient suggests that highly qualified responders take less sick leave per year than other employees, perhaps because they have a higher perceived level of importance, as discussed by Rantanen and Tuominen (2011).

Our findings did not support our second hypothesis that employees enrolled in the Federal Employee Retirement System (FERS) plan will have significantly lower absenteeism proportions than those enrolled in the Civil Service Retirement System (CSRS). We hypothesized that employees working under the FERS system may feel incentivized to come to work while sick because that leave time can be applied to early retirement. However, the average amount of sick leave taken for employees under the FERS was 7.3 sick days per year and the average amount of sick leave taken for employees under the CSRS was 7.5 sick days per year. These numbers are negligibly different, and right on par with the population's average proportion of sick leave per year. Several studies, such as those by Aronsson et al. (2000), Caverly et al. (2007), and Dew et al. (2004), have illustrated the negative impact sick presenteeism can have on the productivity and overall functioning of a workforce. It is heartening to conclude that there is no evidence of employees on FERS feeling incentivized to save their sick leave, which has been linked to sick presenteeism.

The final predictive model we developed of sick absenteeism at the CDC incorporated age, gender, response Tier and an interaction term between age and gender. This final model was quite different than the one proposed in our third hypothesis, in which employee GS level, length of service at the CDC and response Tier qualification were the predicted parameters. The negative linear association between response Tier qualification and yearly sick leave proportions remained after controlling for age and gender, though the average proportion of sick leave per year increased slightly across all Tier levels.

Unfortunately, our modeling capabilities were limited for some of the interaction variables due to their sparse distribution. For example, we were unable to assess potential interactions between response Tier and length of service at the CDC or retirement system. An elaboration and further explanation of this limitation can be found in the strengths and limitations section below. Despite our limited assessments, our final model did include an interaction term between age and gender. Martocchio (1989) observed a strong trend of declining proportions of absenteeism among older men, and a similar trend among older women. Our model does not support his observations. The younger age groups for both men and women in our study population had lower proportions of sick absenteeism than the older age groups, with the exception of women between the ages of 45-54. This model also contradicts Vandenheuvel and Wooden (1995), which concluded that the youngest and oldest age groups had the highest proportions of absenteeism.

Strengths and Limitations

All attendance data were collected automatically by an automated electronic system in fifteen minute increments. While time is technically self-reported, that reporting undergoes considerable oversight by supervisors, and is as accurately captured as possible. Since it is electronically recorded from the beginning, there is considerably less of a chance that it could be transcribed incorrectly or was assembled into our database inaccurately. Since our demographic data were linked from employment records and not reliant on self-reporting, it is also as accurate as possible.

Due to privacy concerns, many of our variables were obtained as categorical values instead of continuous values. While the categories were relatively small, it is possible that some variation in sick leave was lost or obscured. We attempted to minimize lost variation by requesting categories that reflected natural divisions in the variable; for example, the first category in the length of service variable was 0-4 years and the next was 5-9 years because benefits change slightly for employees at year five and year ten. We also ended up collapsing some of the categories for variables such as age and pay grade because they were highly similar. For example, the average amount of sick leave for those aged 35-39 years was 7.6 sick days per year and those aged 40-44 years was 7.7

sick days per year, therefore they were combined into one category. The high similarity between some of our categories suggests that any loss in variation when the continuous variables were converted into categorical variables was minimal.

Unfortunately, the categorical nature of several variables limited us to recording a single value for each participant. The values we selected were the ones at the end of the study timeframe, or end of that participant's work at the CDC. For example, if an employee started working in 2004 and was still working at the end of our observation period in 2012, we recorded their age as of 2012. This may have resulted in some partially misleading categorizations, for example if an employee's fifth year at the CDC was in 2012, they would be counted in the 5-9 year category even though for a majority of their time in this study they worked less than 4 years. In future studies this issue could be resolved if data was delivered on every individual for every year they worked during the study timeframe. That way trends could be assessed on a yearly basis and individuals could switch categories during the timeframe.

Our conclusions about the length of service at the CDC variable may be susceptible to error due to our variable definition. We only examined the length of time an employee worked at the CDC, without considering their history of other federal employment. It is possible that some employees have extensive federal experience but minimal time at CDC, but only their time at CDC would be captured in our analysis. Therefore, if someone had a total of over fifteen years of federal service, but only 3 years of service at the CDC, they would be treated as if they only had 3 years of service in total. This assumption could cause some confounding in relation to other variables, such as age and pay grade. For example, there may be an outlier where someone with only three years of experience is working at the highest possible pay grade and is 55 years old, whereas the majority of people with three years of experience are younger and have lower pay grades. Unfortunately, it was impossible to use total length of federal service for our analyses instead of length of service at the CDC because response Tier qualification is linked almost entirely to the amount of time someone works at the CDC specifically.

Determining how to use the sick leave variable in our analysis was a considerable challenge. Since not all of the participants worked during the entire study timeframe, it was not possible to compare the overall amount of sick leave across the board. In addition, many participants did not work the maximum number of possible hours each year, which made proportional assessments impossible. Not only were the proportional estimates extremely small (because most employees take relatively few sick days compared to days worked), but they also were influenced by other types of leave that employees took, which was not a focus of this study. Therefore, we ultimately decided to convert sick leave into a proportion for each observation. This proportion assumes a standardized number of 250 work days per year, which is not entirely accurate. However, we felt that there was not a need to get more specific, and that this number would eliminate unnecessary work while still yielding nearly the same results as a more accurate figure.

Due to the extremely large size of our dataset, we needed to apply nonconventional assessments of statistical significance to our results. All of our initial statistical tests yielded results that were significant at an alpha level of 0.001. However, these values were not necessarily meaningfully significant in terms of actual differences in sick leave proportions. After consulting with my supervisor at the CDC, we determined that a meaningful difference in sick leave use proportions had to be greater than one day (8 hours) per year. This cut-point for meaningful significance was applied to all of our further analyses. While we acknowledge that this cutoff point for significance is somewhat arbitrary, it is no more arbitrary than using alpha significance values. Our preliminary model assessments ruled out non-significant variables by using an alpha value of 0.0001, which we found was roughly correlated to a difference of 1 day per year, and then we assessed the remaining variables by comparing their parameter estimates to our 1 day per year rule to ensure meaningful significance.

There were a few potential interaction terms that we were not able to assess while creating a model due to small cell sizes. Three of these interaction terms were associated with response Tier qualification. These terms, response Tier*pay grade, response Tier*length of service, and response Tier*retirement plan, all contained one or more zero cells. In addition, there were two age-related interaction terms that contained zero cells, age*length of service at the CDC and age*retirement system. None of these five terms were included in any potential models because their zero cells were could not be interpreted by the statistical tests we were running. It is unfortunate that we were unable to assess the potential impact of these interaction terms on a model and future studies should structure the data in a way that allows for their analysis.

Our analysis was limited to single or two-way interactions, however it may be worthwhile for future researchers examine potential three-way interactions between these demographic variables. For example, there may be an interesting relationship between gender, age, and response Tier qualification. Conducting such analysis was beyond the scope of this study, and would require extremely complex modeling techniques.

Finally, our analysis only examined sick leave proportions within this population. An analysis of the demographic trends behind total leave proportions might have considerable value to internal management and planning staff.

Conclusion

Sick absenteeism proportions were observed to vary significantly (variation of greater than 1 sick day per year) by gender, EOC response Tier, length of service at the CDC, age, and GS pay grade level. These findings support our first hypothesis that sick leave proportions will vary significantly by five of our six key variables of interest.

Women took on average two full days of sick leave per year more than men. This finding echoes the conclusions drawn by Leigh (1983) and Meisenheimer (1990) that women tend to take significantly more sick leave per year than men. Surprisingly, the youngest age group in our study (20-24 years old) took almost a day less of sick leave per year than the oldest age group (55+ years). Participants between 35-44 years old took the most sick leave of any age group--a n average of one day per year. This finding, as well as our findings on the association between pay grade and absenteeism, contradict previous literature and merit further investigation in this population.

There were some concerns that the new FERS retirement system incentivized sick presenteeism; however, we did not find any evidence of this phenomenon during our analysis. Interestingly, we did observe a linear relationship between response Tier qualification and sick leave absenteeism proportions. Those among the higher response Tiers took significantly less sick leave than those among the lower response Tiers, and this trend remained after controlling for age and gender.

This study will provide invaluable information to CDC management and planning officials on sick leave proportions amongst the workforce. It was the first of its kind to examine the relationship between demographics and absenteeism at the CDC, and provides an initial stepping stone for further investigation into these complex associations. Future studies should examine these associations on smaller time scales, perhaps breaking the data down by month or even day of the week. A smaller timescale would allow investigators to examine the monthly trends in sick leave, which could reveal trends related to flu season or school holidays such as spring break. Leonard et al. (1990) conducted a five-year time series analysis predicting total absence frequency by a month-season-year model, and concluded absenteeism peaked in the winter about the same time as the annual flu. Linking demographic variables with sick absenteeism may highlight groups within the workforce in need of interventions, such as receiving a yearly flu vaccine. The CDC could utilize that information to tailor infection control or flu shot messaging to their employees.

Evaluating and understanding trends in sick and total absenteeism use provides a plethora of data for organizations to utilize while developing their strategic policies, such as continuity of operations and capacity loss mitigation. Linking absenteeism to demographic variables could help organizations tailor presenteeism programs or campaigns to unique subpopulations within their workforce, and ultimately may promote better attendance.

References

Aronsson, G., Gustafsson, K. & Dallner, M. Sick but Yet at Work: an Empirical Study of Sicknes Presenteeism. Journal of Epidemiology Community Health 2000; 54:502–9.

Bedian, A., Ferris, G., & Kacmar, M. Age, Tenure, and Job Satisfaction: A Tale of Two Perspectives. Journal of Vocational Behavior 1992;40:33-48.

Bungum, T., Satterwhite, M., Jackson, A., & Morrow, J. The relationship of body mass index, medical costs, and job absenteeism. American Journal of Health Behavior 2003; 27(4):456-462.

Camp, S. & Lambert, E.G. The Influence of Organizational Incentives on Absenteeism: Sick Leave Use Among Correctional Workers. Journal of Criminal Justice 2005; 33(2):165-175

Caverly, N., Cunningham, J.B., MacGregor, J.N. Sickness Presenteeism, Sickness Absenteeism, and Health Following Restructuring in a Public Service Organization. Journal of Management Studies 2007; 44(2):304-319.

De Boer, E.M., Bakker, A., Syroit, J.E., & Schaufeli, W.B. Unfairness at Work as a Predictor of Absenteeism. Journal of Organizational Behavior 2002; 23(2):181-197.

Dew, K., Keefe, V., & Small, K. Choosing to Work When Sick: Workplace Presenteeism. Social Science & Medicine 2004; 60(2005):2273-2282.

Farrell, D. H. Commitment, Absenteeism and Turnover of New Employees: A Longitudinal Study. Human Relations 1984;37:681-692.

Goetzel, R. Z., Hawkins, K., Ozminkowski, R., & Wang, S. An Analysis of the Health and Productivity Cost Burden of the Physical and Mental Health Conditions Affecting Six Large Corporations in 1999. Value in Health 2003; 6(3):211-211.

Goetzel, R. Z., Long, S. R., Ozminkowski, R., Hawkins, K, & Wang, S. Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting US employers. Journal of Occupational and Environmental Medicine 2004; 46(4):398-412.

Goldberg, C.B., & Waldman, D. Modeling Absenteeism: Testing Alternative Measures and Mediated Effects Based on Job Satisfaction. Journal of Organizational Behavior 2000;21(6):665-676.

Grinyer, A., & Singleton, V. Sickness Absence as Risk Taking Behaviour: A Study of Organisational and Cultural Factors in the Public Sector. Health, Risk and Society 2000; 2:7–21.

Harrison, D. A. & Martocchio, J. J. Time for absenteeism: A 20-year review of origins, offshoots, and outcomes. Journal of Management 1998;24(3):305-350.

Hemp, P. Presenteeism: at work – but out of it. Harvard Business Review 2004; October: 49–58.

Hendrix, W. H. & Spencer, B. A. Development and Test of a Multivariate Model of Absenteeism. Psychological Reports 1983;64(3):923-938.

Hilton, M. F., Sheridan, J., Cleary, C., Whiteford, H.A. Employee Absenteeism Measures Reflecting Current Work Practices May be Instrumental in a Re-evaluation of the Relationship Between Psychological Distress/Mental Health and Absenteeism. International Journal of Methods in Psychiatric Research 2009; 18(1):37-47.

Knutsson, A. & Goine, H. Occupation and unemployment rates as predictors of long term sickness absence in two Swedish counties. Social Science Medicine 1998; 47:25–31.

Lambert, E. G., Camp, S. D., Edwards, C., & Saylor, W. G. Here Today, Gone Tomorrow, Back Again the Next Day: Absenteeism and Its Antecedents Among Federal Correctional Staff. Journal of Criminal Justice 2005; 33(2):165-175.

Leigh, J. P. Sex-Differences in Absenteeism. Industrial Relations 1983; 22(3):349-361.

Leonard, C., Dolan, S. L., Arsenault, A. Longitudinal Examination of the Stability and Variability of 2 Common Measures of Absence. Journal of Occupational Psychology 1990;63(4):309-316.

Leonard, C., Vanameringen, M. R., Dolan, S., & Arsenault, A. Absenteeism and Attendance at Work. Relations Industrielles-Industrial Relations 1987; 42(4):774-789.

Lowe, G. Here in Body, Absent in Productivity: Presenteeism Hurts Output, Quality of Work-life and Employee Health. Canadian HR Reporter: The National Journal of Human Resource Management 2002; 16:1–2.

Lu, X. S., Laaksonen, M., Aittomaki, A., Leino,Arjas, P., Rahkonen, O., Saatamoinen, P., & Lahelma, E. Sickness Absence, Employment History, and High-Risk

Employees: A 10-Year Longitudinal Study. Journal of Occupational and Environmental Medicine 2010; 52(9):913-919.

Martocchio, J. J. Age-Related Differences in Employee Absenteeism - a Meta-Analysis. Psychology and Aging 1989; 4(4):409-414.

Martocchio, J. J. The Effects of Absence Culture on Individual Absence. Human Relations 1994; 47:243-262.

Mason CM, Griffin MA. Group absenteeism and positive affective tone: a longitudinal study. J Organ Behav 2003; 24: 667–687.

McKevitt, C., Morgan, M., Dundas, R., & Holland, W. Sickness absence and 'working through' illness: a comparison of two professional groups. Journal of Public Health Medicine 1997; 19(3):295–300.

Meisenheimer, J. R. Employee Absences in 1989 - a New Look at Data from the Cps. Monthly Labor Review 1990; 113(8):28-33.

Oshagbemi, T. Is Length of Service Related to the Level of Job Satisfaction? International Journal of Social Economics 2000; 27(3):213-226.

Office of Personnel Management. (2012). Pay and Leave. https://www.opm.gov/policy-data-oversight/pay-leave/

Rantanen, I., & Tuominen, R. Relative Magnitude of Presenteeism and Absenteeism and Work-related Factors Affecting Them Among Health Care Professionals. International Archives of Occupational and Environmental Health 2011; 84(2):225-230.

Rentsch, J., Steel, R. Testing the Durability of Job Characteristics as Predictors of Absenteeism Over a Six-Year Period. Personnel Psychology 1998; 51:165-190.

Sagie, A. Employee Absenteeism, Organizational Commitment, and Job Satisfaction: Another Look. Journal of Vocational Behavior 1998; 52:156-171.

Saltzstein, A.L., Ting, Y., & Saltzstein, G.H. Work-family Balance and Job Satisfaction: The Impact of Family-Friendly Policies on Attitudes of Federal Government Employees. Public Administration Review 2001; 61(4):452-467.

Schappi, J. V. (1988). Improving Job Attendance. Washington, D.C.: The Bureau of NationalAffairs.

Scott, K. D., Markham, S. E., & Robers, R. W. Rewarding Good Attendance: A Comparative Study of Positive Ways to Reduce Absenteeism. Personnel Administrator 1985; 30:72-83.

Sharp, C. & Watt, S. A Study of Absence Rates in Male and Female Employees Working in Occupations of Equal Status. Occupational Medicine-Oxford 1995; 45(3):131-136.

Spears, D.R., McNeil, C., Warnock, E., Trapp, J., Oyinloye, O., Whitehurst, V., Decker, K.C., Chapman, S., Campbell, M., Meechan, P. Predicting Temporal Trends in Sickness Absence Rates for Civil Service Employees of a Federal Public Health Agency. J. Occupational & Environmental Medicine 2013; 55(2):179-190.

van Vuuren, M., de Jong, M. D. T., & Erwin, R.S. Commitment With or Without a Stick of Paid Work: Comparison of Paid and Unpaid Workers in a Nonprofit

Organization. European Journal of Work and Organizational Psychology 2008; 17(3):315-326.

Vandenheuvel, A., & Wooden, M. Do Explanations of Absenteeism Differ for Men and Women. Human Relations 1995; 48(11):1309-1329.

Wilson, N., & Peel, M.J. The Impact on Absenteeism and Quits of Profit-Sharing and Other Forms of Employee Participation. Industrial & Labor Relations Review 1991; 44(3):454-468.

Williams, M. L., & MacDermid, S. M. Linkages between Employee Benefits and Attitudinal and Behavioral Outcomes: A Research Review and Agenda. Human Resource Review 1994; 4:131-160.

Tables

Table 1. Descriptive statistics for full-time civil servants working at the CDC in Atlanta, GA between 2004-2012

			Sick Leave				
	N	%	Mean*	Median*	Q1	Q3	P-value**
Gender							
Male	2407	35.5	6.1	5.5	2.5	9.1	<0.0001
Female	4374	64.5	8.0	8.1	4.7	10.9	
Age+							
20-34	576	8.5	6.5	5.9	3.0	9.5	<0.0001
35-44	1627	23.9	7.7	7.7	4.2	10.5	
45-54	2356	34.7	7.4	7.2	3.9	10.5	
55+	2222	32.7	7.1	7.0	3.4	10.6	
Length of Service+							
0-3	1607	23.7	5.9	5.2	2.4	8.7	<0.0001
4-9	1572	23.1	7.1	7.1	3.6	10.1	
10-19	2210	32.5	8.0	8.3	4.8	11.1	
20-29	1229	18.1	8.1	8.1	5.0	11.2	
30+	163	2.4	7.0	7.4	2.0	10.8	
Responder Tier							
0	4917	72.5	7.5	7.3	3.8	10.6	<0.0001
1	750	11.1	7.5	7.8	4.4	10.6	
2	657	9.7	6.8	6.6	3.4	10.0	
3	339	5.0	6.1	5.7	2.8	9.1	
4	118	1.7	5.3	4.1	2.1	8.2	
Pay Grade							
01-03	253	3.7	4.3	3.4	1.2	6.6	<0.0001
04-07	275	4.1	10.2	10.8	7.8	12.1	
08-15	6253	92.2	7.3	7.2	3.8	10.4	
Retirement System							
CSRS	423	6.2	7.5	7.8	3.2	11.2	0.155
FERS	6358	93.8	7.3	7.2	3.7	10.5	

** sick days/year*

***Type III Analysis*

+in years

Table 2. Parameter estimates and type III analysis results for sick leave use per year of civil servants at the CDC in Atlanta, GA between 2004-2012 by four predictor variables

					Type III Analysis	
		Parameter estimate	Difference*	Proportion of Sick Leave**	DF	P-value
Age+					3	<0.0001
	20-34	-0.0018	1.00	6.5		
	35-44	0.0994	1.10	8.6		
	45-54	-0.0192	0.98	6.5		
	55+	*Reference*				
Gender					1	<0.0001
	Female	0.2644	1.30	8.8		
	Male	*Reference*				
Response Tier					4	<0.0001
	0	0.2501	1.28	8.8		
	1	0.2449	1.28	8.8		
	2	0.1561	1.17	8.6		
	3	0.0659	1.07	8.5		
	4	*Reference*				
Age*Gender					3	<0.0001
20-34	Female	-0.0040	1.00	6.5		
20-34	Male	*Reference*				
35-44	Female	-0.0001	1.00	6.5		
35-44	Male	*Reference*				
45-54	Female	0.0630	1.07	8.6		
45-54	Male	*Reference*				
55+	Female	*Reference*				
55+	Male	*Reference*				

** Units are the absolute difference in sick days/year, calculated by exponentiating the parameter estimate, significant if greater than or equal to 1*

*** days per year*

+*in years*

Figures

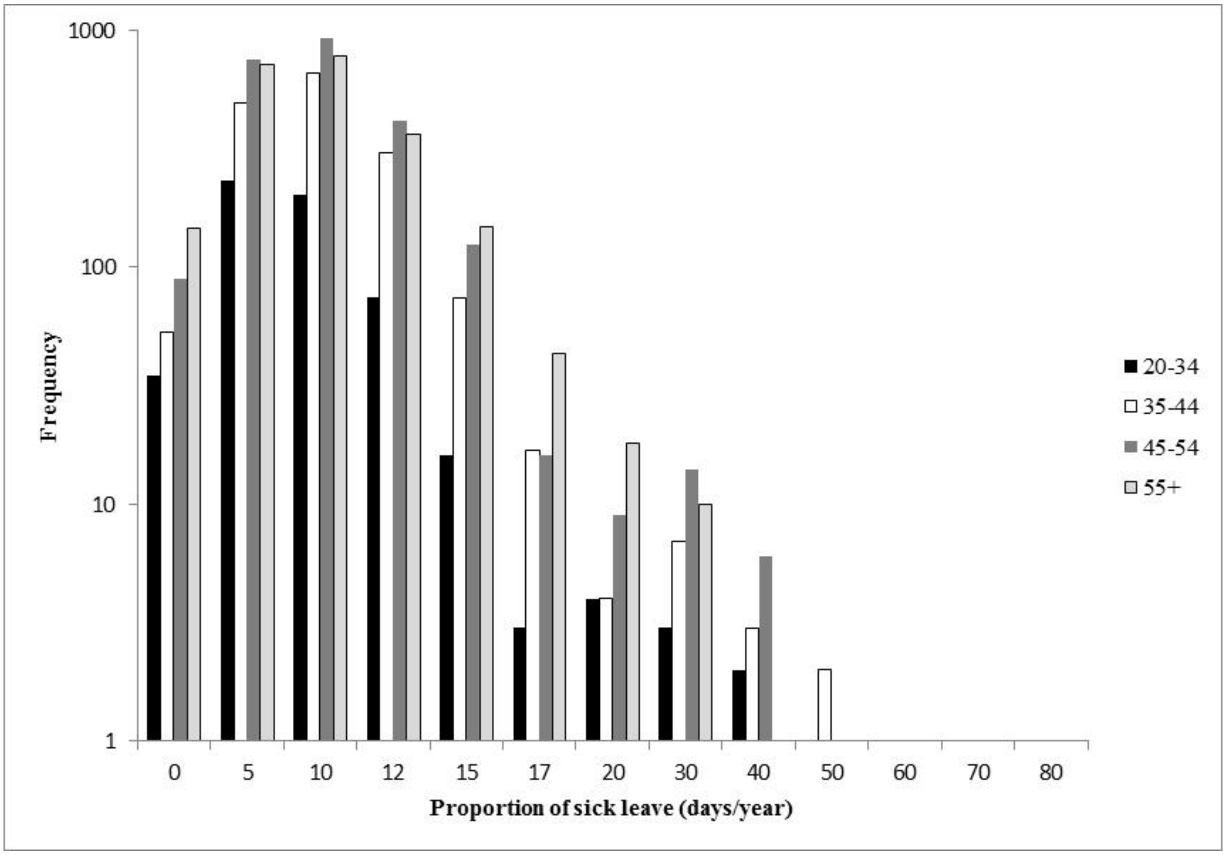

Figure 1. Distribution of sick leave use (days/year)[1] by age group (in years) among full-time civil servants working at the CDC in Atlanta, GA between 2004-2012.

[1] Proportions on X axis are grouped, value "0" represents 0-4.9 sick days per year, "5" represents 5-9.9 sick days per year, etc…

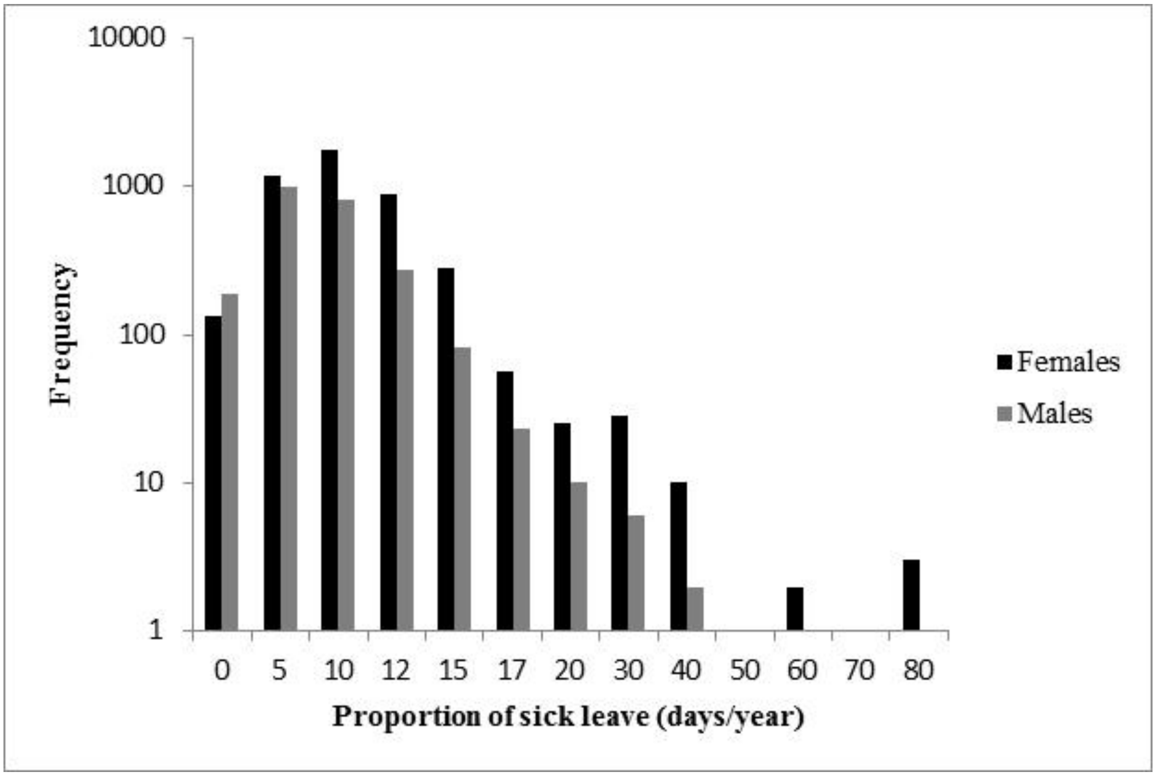

Figure 2. Distribution of sick leave use (days/year)[2] by gender among full-time civil servants working at the CDC in Atlanta, GA between 2004-2012.

[2] Proportions on X axis are grouped, value "0" represents 0-4.9 sick days per year, "5" represents 5-9.9 sick days per year, etc…

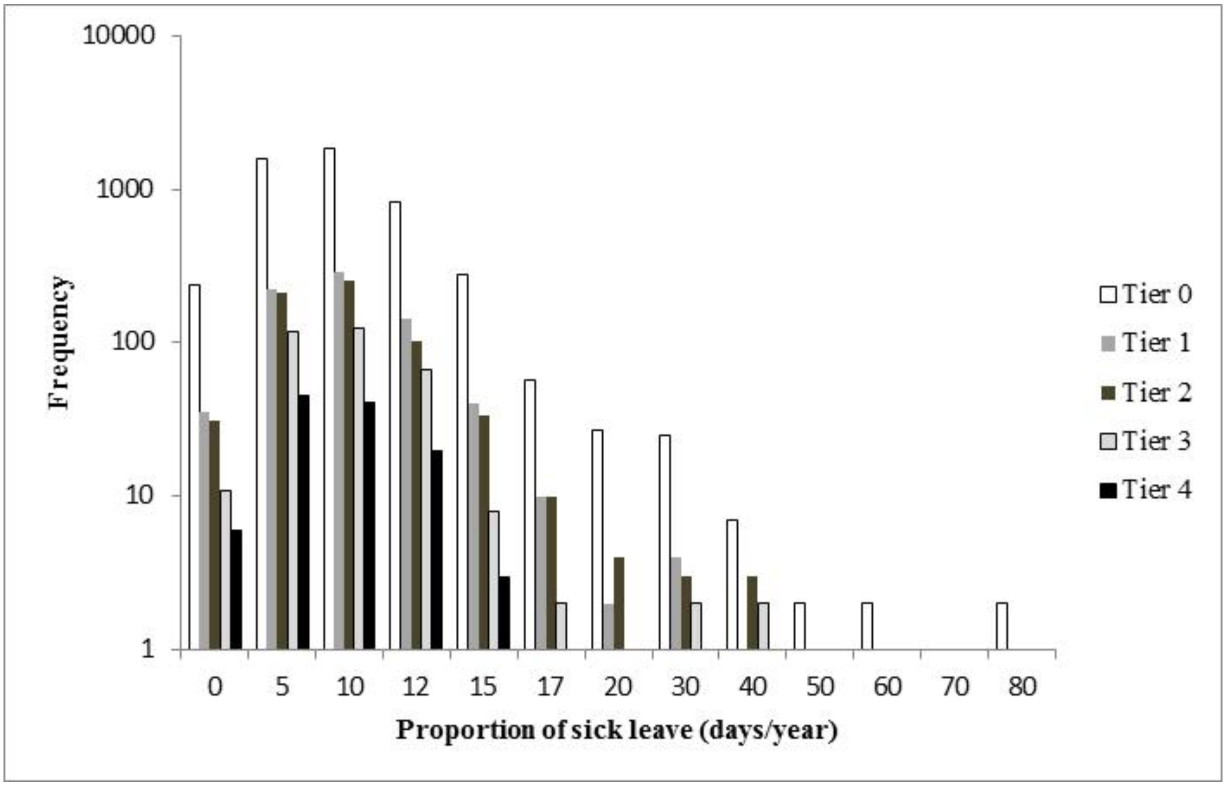

Figure 3. Distribution of sick leave use (days/year)[3] by CDC response Tier qualification among full-time civil servants working at the CDC in Atlanta, GA between 2004-2012.

[3] Proportions on X axis are grouped, value "0" represents 0-4.9 sick days per year, "5" represents 5-9.9 sick days per year, etc…

IV. Summary and future directions

For organizations such as the CDC, understanding trends and variations in absence rates for subpopulations within the workforce is essential for the development of surge capacity plans and targeted leave messaging, which may ultimately contribute to a higher attendance rate and healthier workforce.

Sick absenteeism proportions were observed to vary significantly (variation of greater than 1 sick day per year) by gender, EOC response Tier, length of service at the CDC, age, and GS pay grade level. Women took on average 2 full days of sick leave per year more than men. This finding echoes the conclusions drawn by Leigh (1983) and Meisenheimer (1990) that women tend to take significantly more sick leave per year than men. Surprisingly, the youngest age group in our study (20-24 years old) took almost a day less of sick leave per year than the oldest age group (55+ years). Participants between 35-44 years old took the most sick leave of any age group, an average of one day per year. This finding, as well as our findings on the association between pay grade and absenteeism, contradicts previous literature and should be investigated further in this population.

People working in healthcare services have been shown to have an increased likelihood of going to work while sick (Rantanen and Tuominen 2011). The most common barriers to taking sick leave among this demographic include: difficulty in replacement, amount of make-up work that must be covered after the absence, and attitudes about their own level of health. However, attendance at work while being ill could have negative impacts on productivity and may ultimately lead to greater proportions of absenteeism in that office, if other employees contract the illness. There were some concerns that the new FERS retirement system incentivized sick presenteeism, however we did not find any evidence of this phenomenon during our analysis. From a public health perspective, this is an extremely positive finding. The CDC depends on its staff to protect the health of the United States and to respond rapidly and fully to any public health emergencies. Therefore, any policy that may encourage sick presenteeism should be changed immedieatly.

Interestingly, we observed a linear relationship between response Tier qualification and sick leave absenteeism proportions. Those among the higher response Tiers took significantly less sick leave than those among the lower response Tiers. Further, the proportion of sick leave taken by those among the highest response Tier was significantly lower than the population's average proportion of sick leave (5.3 days per year vs. 7.3 days per year). This finding may reflect attitudes of low replaceability among upper level response staff, and warrants further investigation because it may also reflect greater sick presenteeism trends in those staff. High proportions of sick presenteeism among upper level response staff could become a serious issue if it impacts productivity or results in larger numbers of employees falling ill in the same timeframe. Despite the presence of hand sanitization stations throughout the Emergency Operations Center (EOC), there is a significant risk of spreading an infectious disease to other responders during a public health response. Responders work in incredibly close quarters for very long hours and typically share office equipment, such as computer stations and phones, with multiple shifts of people. If response leads are more likely to come to the EOC with an infectious disease that is transmitted via air or droplets, such as influenza, they could infect many more response staff and hinder the overall response capacity of the CDC. Further investigations, perhaps through qualitative surveys, should be conducted to determine why proportions of sick leave are significantly lower amongst upper level responders. If those investigations conclude that these responders have a high proportion of sick presenteeism, then new policies or campaigns may have to be initiated specifically for this population, deterring them from attending work while ill.

This study will provide invaluable information to CDC management and planning officials on sick leave proportions amongst the workforce. It was the first of its kind to examine the relationship between demographics and absenteeism at the CDC, and provides an initial stepping stone for further investigation into these complex associations. Future studies should examine these associations on smaller time scales, perhaps breaking the data down by month or even day of the week. A smaller timescale would allow investigators to examine the monthly trends in sick leave, which could reveal trends related to flu season or school holidays such as spring break. Leonard et al. (1990) conducted a five-year time series analysis predicting total absence frequency by a month-season-year model, and concluded absenteeism peaked in the winter about the same time as the annual flu. Linking demographic variables with sick absenteeism may highlight groups within the workforce that perhaps are not receiving a yearly flu vaccine. The CDC could utilize that information to tailor infection control or flu shot messaging to their employees.

Future studies on the relationship between demographic variables and total absenteeism trends should also be conducted in this population. While sick leave is of particular interest for internal operations and response purposes, total leave would provide a much clearer picture of exactly how many civil servants are at work and what those workers look like.

Appendices

Appendix A

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control
and Prevention (CDC)

Memorandum

Date	October 11, 2012
From	Barbara R. DeCausey, MPH, MBA Chief, Human Research Protection Office
Subject	HRPO Exemption Determination for Protocol #6362.0, "Temporal Trends of Absenteeism and Pertinent Implications for Internal Emergency Response Planning"
To	Kimberly Gajewski OCOO/OSHE

On behalf of the CDC Human Research Protection Office (HRPO), I have reviewed the request to exempt protocol #6362.0, "Temporal Trends of Absenteeism and Pertinent Implications for Internal Emergency Response Planning" and find that this research activity is exempt under 45 CFR 46.101(b)(2). This determination is valid for a period of three years through **10/10/2015.** However, we strongly encourage investigators to close out exempt protocols as soon as CDC staff are no longer engaged in the research activity, rather than waiting for a reminder of the three-year expiration date.

<u>Please be aware that changes to this protocol may not be implemented until they are reviewed by HRPO and determined to be consistent with the exemption categories.</u> You will be reminded in three years (if the study has not been completed and closed) to submit another request for continuation and to confirm that no changes have been made to the protocol or the related science that would affect the ethical appropriateness of the research or this exemption determination.

Please also be advised that investigators remain responsible for the ethical conduct of this study and for ensuring appropriate human research protections even for research that is exempt from the regulations governing the protection of human subjects in research.

If you have questions, please contact your Associate Director for Science, your National Center Human Subjects Contact, or HRPO at huma@cdc.gov, or by telephone at 404-639-4721.

cc:
Ross Spears
Carrie McNeil
Natalie Brown

APPENDIX B

Descriptive Statistics using original variable categorization for civil servants working at the CDC between 2004-2012

			Sick Leave			
	N	%	Mean*	Median*	Q1	Q3
Gender						
Male	2407	35.5	6.1	5.5	2.5	9.1
Female	4374	64.5	8.0	8.1	4.7	10.9
Age						
20-24	2	0.0	5.7	5.7	5.2	6.2
25-29	116	1.7	5.6	4.7	1.9	8.1
30-34	458	6.8	6.7	6.3	3.2	9.9
35-39	732	10.8	7.6	7.6	3.9	10.4
40-44	895	13.2	7.7	7.8	4.5	10.7
45-49	1122	16.6	7.4	7.2	3.9	10.4
50-54	1234	18.2	7.4	7.3	4.0	10.5
55-59	1079	15.9	7.1	7.1	3.6	10.5
60+	1143	16.9	7.2	7.0	3.2	10.9
Length of Service						
0-3	1607	23.7	5.9	5.2	2.5	8.6
4-9	1572	23.2	7.1	7.1	3.6	10.1
10-14	1601	23.6	8.0	8.3	4.8	11.1
15-19	609	10.0	8.0	8.4	4.9	11.2
20-24	885	13.1	8.0	8.1	5.0	11.1
25-29	344	5.1	8.4	8.4	4.6	11.3
30+	163	2.4	7.0	7.5	2.0	10.7
Responder Tier						
0	4917	72.5	7.5	7.3	3.8	10.6
1	750	11.1	7.5	7.8	4.4	10.6
2	657	9.7	6.8	6.6	3.4	10.0
3	339	5.0	6.1	5.7	2.8	9.1
4	118	1.7	5.3	4.1	2.1	8.2
Pay Grade						
00-03	253	3.7	4.3	3.4	1.3	6.6
04-07	275	4.1	10.2	10.8	7.8	12.1
08-11	917	13.5	8.7	9.1	5.3	11.4
12-15	5336	78.7	7.8	6.9	3.5	10.1
Retirement System						
CSRS	423	6.2	7.5	7.8	3.2	11.2
FERS	6358	93.8	7.3	7.2	3.7	10.5

**sick leave days/year*

APPENDIX C

Initial significance test of all possible interaction terms*					
	Responder Tier	**Age**	**Sex**	**Length of service**	**Retirement system**
Pay Grade	0.13	<0.0001	0.002	<0.0001	0.04
Responder Tier		0.84	0.58	0.004	<0.0001
Age			<0.0001	<0.0001	0.1251
Sex				0.08	0.023
Length of Service					<0.0001

**significance when p<0.0001*

APPENDIX D

```
*********Initial test of the full model;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire/dist=poisson type3;
run;
quit;

***Tesing interactions, one at a time, starting with pay grade;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire grade*retire/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire grade*tier/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire grade*age/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire grade*sex/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
```

```
class retire;
model bestrsick= grade tier age sex CDC retire grade*CDC/dist=poisson type3;
run;
quit;

****testing response tier interactions;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire tier*age/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire tier*sex/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire tier*CDC/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire tier*retire/dist=poisson type3;
run;
quit;

****Testing age interations;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire age*sex/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
```

```
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire age*CDC/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire age*retire/dist=poisson type3;
run;
quit;

***Testing gender interactions;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire sex*CDC/dist=poisson type3;
run;
quit;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire sex*retire/dist=poisson type3;
run;
quit;

****Testing length of service interactions;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= grade tier age sex CDC retire CDC*retire/dist=poisson type3;
run;
quit;

****full model, dropped interaction terms that were not individually
significant;
** Also dropped three interaction terms that had small cell sizes <10;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= tier age sex CDC retire grade grade*age grade*cdc
age*sex/dist=poisson type3;
run;
```

```
quit;

***dropping grade*CDC;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= tier age sex retire CDC grade grade*age age*sex /dist=poisson
type3;
run;
quit;

*** dropping age*grade;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
class retire;
model bestrsick= tier age sex retire CDC grade age*sex /dist=poisson type3;
run;
quit;

***dropping retirement;
proc genmod data=thesis2;
class grade;
class tier;
class age;
class CDC;
class sex;
model bestrsick= tier age sex CDC grade age*sex /dist=poisson type3;
run;
quit;

**dropping grade;
proc genmod data=thesis2;
class tier;
class age;
class CDC;
class sex;
model bestrsick= tier age sex CDC age*sex /dist=poisson type3;
run;
quit;

**dropping CDC, this is the final model;
proc genmod data=thesis2;
class tier;
class age;
class CDC;
class sex;
model bestrsick= tier age sex age*sex /dist=poisson type3;
run;
quit;
```