

Employee Assistance Program Counseling Improves Clinical and Work Outcomes: Longitudinal Results from CuraLinc Healthcare 2017-2022 in United States

Dr. Mark Attridge*, David Pawlowski**, and Sean Fogarty**

* Attridge Consulting, Inc., United States

** CuraLinc, LLC (d.b.a. CuraLinc Healthcare), United States

DOI: 10.29322/IJSRP.12.08.2022.p12802

<http://dx.doi.org/10.29322/IJSRP.12.08.2022.p12802>

Paper Received Date: 13th July 2022

Paper Acceptance Date: 29th July 2022

Paper Publication Date: 6th August 2022

Abstract: This paper presents empirical findings from a multi-year applied naturalistic study that focused on changes in clinical and work outcomes after using an employee assistance program in the United States. Self-report outcomes assessed with standardized measures were collected at the start of counseling from 33,683 employees during the normal course of business at CuraLinc Healthcare during 2017 to June 2022. A total of 5,159 employees had valid Pre and 30-day Post use data on one or more of the outcomes. Tests within each outcome sample found significant improvement after counseling (most had a 5 or 6 session model). Among the 487 cases using the EAP for a depression issue, the average severity level of depression symptoms (PHQ-9) was reduced by 59%. Among the subset of these cases at-risk for clinical depression disorder ($n = 292$), 89% had a reliable clinical improvement in severity. Among the 440 cases using the EAP for an alcohol issue, the average severity level of alcohol misuse (AUDIT-10) was reduced by 67%. Among subset of these cases at-risk for alcohol misuse disorder ($n = 288$), 73% had a reliable clinical improvement in severity. Across all cases with absence data ($n = 3,732$), the average hours of work absence in the past month per employee (Workplace Outcome Suite) was reduced by 80%, changing from 8.9 hours at Pre to 1.8 hours at Post. At the start of counseling, 35% of these cases were classified as having an absence problem (i.e., missing more work than a typical healthy employee), but at Post only 7% had a work absence problem. Across all cases with productivity data ($n = 3,845$), the level of work productivity (Stanford Presenteeism Scale) improved by 35%. At the start of counseling, 34% of cases were initially classified as having a problem with work productivity (i.e., low performance and lack of focus), but at Post this rate was reduced to just 5% of cases. All primary results for each outcome were large size statistical effects. Moderator tests indicated the improvement in each outcome was generally consistent across different sub-groups of employees based on client age, sex, clinical use characteristics and other study context factors. Some moderator findings were identified for clinical issue, referral type (self or manager) and client sex having slightly different profiles on certain outcomes. Comparisons with past research, study design limitations, and best practices in conducting applied research on workplace mental health are discussed.

Index Terms: absenteeism, alcoholism, counseling, depression, employee assistance program, productivity

I. INTRODUCTION

The need for mental health support for working adults all over the world is well documented in the literature [1]. For the United States, in normal times, epidemiological data on prevalence rates indicate that about 1 in every 5 (20%) working adults have diagnosable mental health disorders and/or substance abuse problems [2]. This rate is even higher when examining the full population and not just the employed (typically healthier) segment. The 2018 National Survey on Drug Use and Health found that 23% of adults had either a mental illness or a substance use disorder [3]. These kinds of behavioral health issues have been exacerbated since early 2020 during the ongoing COVID-19 global pandemic [4-6].

1.1. Depression and Alcohol Disorders

Depression is a common mental illness that has a negative impact on the way a person thinks, feels, and acts. Depression is also known as major depressive disorder [7]. Although it ranges from mild to more severe levels of impairment, depression can be a serious illness for some of those affected. It causes feelings of sadness and can result in a person losing interest in activities that were once enjoyable. Depression also impacts the ability to function effectively at home and work. It is often experienced together with anxiety disorders, other mental health disorders, and with substance abuse [7]. Depression is also associated with increased suicide risk [8]. Depression is among the most burdensome of behavioral health disorders, giving rise to considerable adverse effects on activities of daily living,

work functioning, and disability [9-11]. Chronic stress and other behavioral health issues are also deeply disruptive for the individuals experiencing them, as well as their friends and family.

Alcohol is the single most used and abused drug in America. According to national survey data [12], in 2015, over 175 million people in the United States reported using alcohol and 1 in every 3 of these people reported binge drinking in the past month (i.e., having 4 or 5 alcoholic drinks in one session). Alcohol misuse includes a spectrum of severity from drinking above recommended limits (called "risky drinking") to severe alcohol dependence. Most people with alcohol misuse are not alcohol dependent (i.e., addicted) and this relatively less severe portion of people account for most of the morbidity and mortality that is attributed to drinking. For example, alcohol misuse contributes to a total of over 88,000 deaths in the United States each year and 1 in 10 deaths among working adults are due to alcohol misuse [13].

1.2. Consequences of Mental Health Conditions

Mental health and alcohol use problems usually do not get better on their own. The consequences of leaving these issues untreated have a far-reaching impact on both physical and emotional fitness [14]. Numerous studies have demonstrated that depression and other mental health conditions are associated with increased overall health care costs [15,16]. There is also substantial evidence that alcohol, depression, and other common behavioral health risk factors are associated with deficits in employee absenteeism and work productivity [17-19]. The greater impact that behavioral health issues have on work presenteeism than on work absenteeism also is a consistent theme in the literature [20-22]. Employers are justifiably concerned about mental health and substance abuse issues both from a personal perspective for their workers and from a cost and productivity standpoint [23-26]. Therefore, it makes sense for employers to try to prevent or reduce the rates of depression, alcohol misuse and other common behavioral health conditions among employees and family members. Employers have a range of behavioral health benefits and tools to choose from to address the rising tide of behavioral health challenges. Sponsoring an employee assistance program (EAP) is a popular response from employers.

1.3. Employee Assistance Programs (EAPs)

According to the Employee Assistance Professionals Association (EAPA) [27], an EAP is:

“a worksite-based program designed to assist (1) work organizations in addressing productivity issues and (2) employee clients in identifying and resolving personal concerns, including, but not limited to, health, marital, family, financial, alcohol, drug, legal, emotional, stress, or other personal issues that may affect job performance.”

According to the most recent national random sample study of employers in the United States in March of 2021 [28], most employers have an EAP - but it varies by size of the company and market sector. In the public sector, all federal employees have an EAP and 78% of all state and local governments offer an EAP to their workers. In the private sector, 83% of large business (500 or more employees), 66% of medium size employers, and 35% of small employers (under 100 employees) sponsor an EAP benefit. In total, over 70.9 million American workers have access to an EAP.

Depression and other mental health problems are common reasons why employees use brief counseling from EAPs [1]. An analysis of over 90,000 EAP cases in the U.S. from year 1999-2010 [29], found only 3.1% of clients were seeking support for an alcohol issue whereas 55% of users had a psychological issue (however, the number specifically with depression was not identified). Another study of a national EAP vendor in Canada [30] examined 9,105 users of counseling from year 2005 to 2007. Their data indicated that 37% of the total sample of EAP counseling cases meet the criteria for having a clinical depression disorder (based on screening at the start of the case). Interestingly, the vast majority of these cases had sought help from the EAP for a reason *other than* depression or anxiety (i.e., 87% of the group with depression disorder and 93% of the other group with only minimal or mild depression severity).

In a more recent example, an industry report examined the mix of issues for over 29,000 total cases contributed by 35 different EAPs during the years 2010 to 2021 [31]. This study found that alcohol or other addiction issues represented just 3% of the total cases, whereas psychological health issues accounted for 30% of the total – with cases specific to depression being 8% of the total. The other two-thirds of cases in this study were spread across categories of personal life stress (29%), marital, dating or family relationships (19%) or work-related issues (19%). Although the number of employees seeking support for alcohol issues is a small part of overall case-mix at most EAPs today, the full-service EAP approach has a long history of understanding how to identify and address employee alcohol problems by using the power of the workplace [32,33]. Identifying and supporting employees with alcohol problems is even specified as one of the seven “core technologies” of the EAP profession [34].

1.4 Research on EAP Cases for Depression or Alcohol Outcomes

Few examples exist in the literature in the past 20 years that have focused specifically on the effectiveness of EAP counseling for employees with depression or alcohol risks [35-46]. Some of this research has involved a mismatch between the client’s clinical issues

in the study sample and the clinical outcomes assessed. More specifically, some studies evaluating the effectiveness of EAP counseling have inappropriately tested for changes in depression or alcohol outcomes among all cases using the EAP for any reason. A logical consequence of having test samples with the majority of cases being irrelevant to clinical depression or to alcohol disorders is finding only modest improvement in depression symptoms and weak or no improvement in levels of alcohol misuse [40-42].

In contrast, other studies evaluating the effectiveness of EAP counseling have used a different approach that focused only on the subsamples of cases selected to be relevant to depression (i.e., scored above the at-risk level on a standardized depression symptoms screener [38,43-46]). Each of these studies found significant and substantial improvement in the severity level of depression symptom after counseling. The statistical effect sizes for the improvement in these five studies that tested only the EAP cases starting treatment at a clinical level of depression ($d = 1.08$ to 1.62) were more than double what was found in the two studies noted earlier that used general samples of EAP users with a range of issues and relevance to depression ($d = 0.45$ [41] and 0.59 [42]). Matching the clinical condition of the user to the outcome is also standard practice for most research studies assessing the effectiveness of computer and phone App machine-based interventions designed for self-support of specific mental health conditions [47-49].

The lesson from the literature is that in order to conduct a fair test of the counseling impact on cases with depression and especially for cases with alcohol problems (as they are more rare), it is important to first identify the subset of EAP users relevant to these specific clinical outcomes. To do this, however, requires a very large sample to start with to provide an opportunity to find enough relevant cases. The normative data on the mix of client issues for most EAPs suggests it takes about 10 cases to find one case who presents with a depression issue and that it takes about 33 cases to find one case who presents with an alcohol issue.

Focusing on cases at the clinical level of behavioral health risks also allows for conducting a more sensitive statistical analysis of the change in clinical symptom severity within each case after EAP treatment. The Reliable Change (RC) index methodology can be used to mathematically test if the extent of clinical change is beyond chance level using the standardized variance of the outcome scale [50-54]. While introduced over 20 years ago, this analytical approach has only recently been applied to study the effectiveness of brief counseling for anxiety and depression cases at EAPs with just three published examples [44-46]. We know of no prior studies using the RC testing approach with alcohol cases at EAPs.

1.5. Research on EAP Cases for Work Outcomes

The typical non-distressed employee misses less than 4 hours per month due to health reasons [see literature review in 55]. In contrast, the typical user of EAP counseling misses about twice as much time from work [22,31,55-58]. Recent studies reveal that by emphasizing the *average* number of absence hours per case, it obscures the fact that the majority of the employees who use EAP counseling actually report zero hours of work absence both at before and after treatment [55-58]. For example, analysis of over 38,000 cases from 47 different EAPs [31] determined that 59% of cases at the start of counseling reported no absence and this group rose to 79% of cases at follow-up. When an absenteeism problem is defined as missing more scheduled work time than the typical employee (i.e., 4 or more hours total in the past month), this same study found that 32% of cases at the start of counseling had a work absence “problem” but that this was reduced to only 15% of cases at follow-up.

In contrast to work absence, the majority of employees who use EAP counseling do typically experience reduced on-the-job productivity (i.e., presenteeism) during the period of acute distress during the month just before seeking counseling [37-40,42,46,55-58]. The same studies all find that work productivity tend to rebound to a more normal level after counseling is completed. For example, the most recent EAP industry-wide study [31] found that over half of EAP cases (56%) had a work presenteeism problem when starting counseling but this was reduced to only 30% of the total cases at the follow-up a few months after treatment. Thus, most EAP cases when starting counseling report having difficulty concentrating at work and doing their job at a normal level of performance even though they may not be missing much (or any) time away from work.

Moderator tests conducted in large samples of cases [31,57,58] consistently have found that employees with different reasons for using the EAP (i.e., clinical issue) tend to have rather similar profiles on work absenteeism and presenteeism/productivity outcomes, although having depression is linked with greater absence and lost productivity. Thus, unlike specific clinical outcomes that are more appropriate to investigate when matched to the client’s goal for therapy or matched to at-risk clinical status (such as alcohol or depression), work-related outcomes are relevant to investigate in all cases regardless of the reason why they used the EAP. Other demographic and clinical use factors also tend to have only small effects as moderators of improvement in work outcomes after use of EAP counseling.

1.6. Methodological Issues in EAP Counseling Effectiveness Research

More than 100 applied research studies have been conducted around the world evaluating the effectiveness of workplace mental health services and EAPs [58]. Over the years, many scholars have critically reviewed this literature on EAP effectiveness [59-73]. Most of the studies have examined the impact of counseling on aspects of clinical functioning and also on work-related outcomes (such as absenteeism and presenteeism /productivity). The consistent conclusion from every one of these reviews is that counseling from EAPs

is generally effective for most users for reducing clinical distress and improving work-related outcomes. Yet, despite these generally supportive findings, these same reviews and also reports by other critics of EAP [74-77] have noted that most of the source studies tend to share practical limitations associated with analyzing data collected from applied business contexts.

Most of the works examining the effectiveness of EAP counseling are white papers, conference presentations, or articles in trade journals rather than articles published in scientific peer-review journals. Most of these studies featured non-experimental, single-group, longitudinal study designs and involved samples of users of EAP counseling who were not randomly selected. Few studies have empirically tested the representativeness of the sub-group of cases with longitudinal data included in the study compared to the much larger caseload at the EAP who did not complete follow-up measures. Less than half of EAPs today even bother to collect longitudinal outcome data from their users, with many opting for a post-only follow-up satisfaction and outcome survey study design [31]. The validity and reliability of the specific outcomes measures used in most studies is also questionable. Despite having multiple research-validated measures available in the public domain for the last 10 to 20 years, a recent study determined that these high-quality tools are not being used by most of the 101 EAP vendors or internal programs surveyed [31]. More specifically, only 35% of EAPs used a research-validated outcome measure for clinical outcomes and only 24% of EAPs used a research-validated measure to assess work-related outcomes. Thus, the level of scientific rigor in many aspects of EAP service evaluation could be enhanced for most providers.

1.7. Project Overview

CuraLinc Healthcare is a global external vendor of EAP services, based in the United States. In business since 2008, it has over 2,900 employer customers that offer the EAP as a benefit to over 4 million employees. This company specializes in delivering transformative mental health care by marrying technology and personalized advocacy to engage, empower and support employees throughout their care journey. The intake clinicians, also referred to as Care Advocates, were all independently licensed, masters or doctorate level educated mental health professionals. During the initial intake assessment, these clinicians were asked to conduct thorough clinical assessments, make expert referrals and collect study outcome data when relevant. The clinicians also provided consultative follow-ups on all EAP cases.

As part of its ongoing business practice, this EAP also routinely collects several kinds of data relevant to assessing the user satisfaction and outcomes of the services. For example, over the past six years of follow-up anonymous survey results have consistently shown a high level of user satisfaction with the EAP counseling service experience. At a 30-day post use follow-up, these cases were asked: "Overall, how would you rate your experience using the program?" and offered response options of: *Excellent, Very Good, Good, Fair* or *Poor*. When combining the top three options, the results have been remarkably consistent each year: Year 2017: 95.0%; Year 2018: 95.1%; Year 2019: 95.1%; Year 2020: 95.3%; Year 2021: 95.0%; Year 2022: 95.1%; Combined: 95.1%.

1.8. Hypotheses and Research Questions

Given the results in past literature on the general effectiveness of EAP counseling, we also expected to find positive changes in both the clinical and work kinds of outcomes after use of counseling in this study. Thus, we proposed the following four research hypotheses:

- *H1*: Among cases who used the EAP for a depression issue, the severity level of depression symptoms will be reduced from Pre to Post use of the service.
- *H2*: Among cases who used the EAP for an alcohol issue, the severity of symptoms of alcohol misuse will be reduced from Pre to Post use of the service.
- *H3*: Among all cases who used the EAP and reported on their work absenteeism, the number of hours of missed work will be reduced from Pre to Post use of the service.
- *H4*: Among all cases who used the EAP and reported on their work performance, the level of performance will be improved from Pre to Post use of the service.

We also wanted to leverage the client background and operational data available in our study to explore the potential for moderators of improvement in the clinical and work outcomes after use of counseling. Thus, we proposed the following research question:

- *RQ1*: Is improvement in clinical and work outcomes from Pre to Post use of the service consistent across different subgroups of cases defined by employee demographic, clinical use, or employer context factors?

II. METHODOLOGY

2.1. Archival Data

We focused only on employee users in this study. Thus, the experiences of spouses, dependents and retirees of covered employees who used the EAP service were excluded. Users were made aware of the service as a benefit open to all covered employees through a variety of digital, interpersonal and workplace promotional practices. There was no direct cost to the employees in this study, as access to the

EAP was sponsored by their employer. Employees participated voluntarily and were not paid for using the services. The study period spanned 63 months, from the start of January of 2017 through the end of March of 2022, based on the start date of program use. The last case included in the study had a Post use data collection date of June 2 of 2022. The year of use was defined by date of when the employee contacted the program and completed the initial intake assessment (2017 to 2022). The case-level raw data was aggregated into one master dataset and analyzed for the present paper. Across the four outcomes examined, we had valid longitudinal data from 5,159 unique employees who worked at over 500 different employers in the United States. Some of this data has been reported on previously in a series of annual reports of preliminary findings [78-81].

2.2. Counseling Intake, Intervention and Follow-up

Employees accessed the EAP in a variety of ways, most commonly by calling the EAP service and talking on the telephone with a licensed mental health professional. After the initial assessment, each employee was provided a referral to engage with a licensed mental health clinician for short-term counseling. Every employee who requested support was referred to a clinician with a specialty that matched their presenting issue or concern who also had confirmed appointment availability. The counseling was most often delivered in either face-to-face sessions at the counselor's office or remotely via live secure video over the Internet. During the initial assessment, the outcome measure(s) were collected, either over the telephone or from a brief online survey. After the counseling treatment phase was completed, a Care Advocate conducted individual follow-ups with cases about 30 days after the last clinical session to assess the employees' clinical progress, use of any recommended additional support services and determine if any other referrals were needed. At this point, the second wave of data was collected for each relevant outcome measure. In all cases, the counselor who treated the employee was not the same person who collected the outcome data. The typical case had a model that allowed for a maximum of 5 or 6 sessions of counseling. The typical cases engaged in treatment over a 50-day period (i.e., from date of case open to case close) with the follow-up survey done at approximately 30 days later after the final session (see Figure 1). Note, the total amount of time involved for each employee to complete their use of the service and participate in the follow-up varied from case to case.

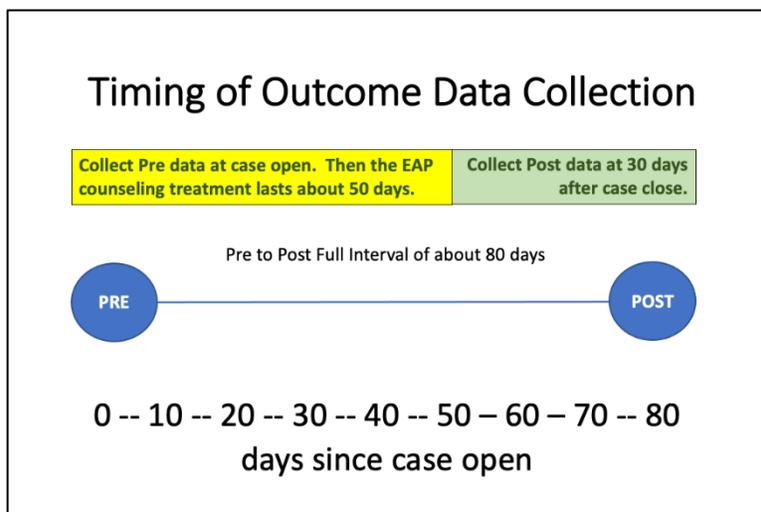


Figure 1: Timing of longitudinal outcome data collection for a typical case.

2.3. Operational System Data

The first kind of data came from the operational business processes used by the staff and counselors who provided the counseling services. Part of this process involves recording core aspects of the business customer context, employee demographics and the clinical experience. For this study we extracted the following information from the operational data system: name of employer/customer, maximum clinical sessions allowed per case in the employer/customer contract, date of first use of the service, date of follow-up survey (only available for cases with the work absenteeism outcome), employee age (date of birth), employee gender, source of referral to the EAP (self or formal referral from management), primary clinical issue (alcohol, depression, work and so on) and the clinical modality.

2.4. Self-Report Outcomes Data

Clinical symptom outcome data was collected at Pre for cases that had a relevant clinical issue (i.e., the case had either depression or alcohol as a primary or secondary issue as the reason for using the EAP). Work outcome data at Pre was collected for cases regardless of the specific clinical issue. Not all relevant EAP cases with depression or alcohol issues were invited to complete the depression or

alcohol clinical symptom measure and not all users completed the work outcome measures. This was because many employees had limited time available at the intake session or were not interested in engaging in the outcome measurement process. The business needs and the unique individual preferences of the employees and/or the intake Care Advocates ultimately determined which specific cases participated in the outcomes study data collection efforts. The Post use data was collected routinely for cases that had the same outcome(s) collected at the start of the program use.

2.5. Outcome Measures

Standardized measures of clinical symptoms and work outcomes were assessed using published and validated scales from the scientific literature and available in the public domain.

Depression. Depression was measured for employees who presented with depression during the initial assessment as either their primary or secondary issue. The Patient Health Questionnaire 9-item scale (PHQ-9) [82] was used for screening, diagnosing, monitoring, and measuring the severity of depression. This scale has been used in many research studies and has established validity and reliability [83-86]. The instructions state: “Over the last 2 weeks, how often have you been bothered by any of the following problems?” It has four response options of: (0) *Not at all*; (1) *Several days*; (2) *More than half the days*; and (3) *Nearly every day*. The PHQ-9 is scored by adding together the scores for all 9 items. Higher scores on this measure indicate greater depression. Scores are categorized into five levels of severity: Minimal = 0 to 4; Mild = 5 to 9; Moderate = 10 to 14; Moderately Severe = 15 to 19; and Severe 20-27. Based on recent reviews [87,88] clinical at-risk status for depression was categorized as Moderate or above (i.e., scores of 10+). This scale had excellent psychometrics with high internal consistency at both time points ($\alpha = .85$ Pre; $.92$ Post) and a significant test-retest correlation ($r_{\text{paired}} = .51$).

Alcohol Misuse. Developed by the World Health Organization, the Alcohol Use Disorders Identification Test, 10-item full scale second version (AUDIT-10) was used to assess level of alcohol misuse and risky drinking behavior. First developed in 1989 for use in a primary care setting, the AUDIT-10 is now used in a wide variety of clinical and research settings as a brief screening tool for identifying at-risk alcohol consumption [89-92]. It measures three content domains: hazardous alcohol use (items 1-3); symptoms of alcohol dependence (items 4-6); and harmful alcohol use (items 7-10). The responses are on a 5-point scale for items 1 through 8 (scored as 0, 1, 2, 3, or 4) and a 3-point scale for items 9 and 10 (scored as 0, 2 or 4). The AUDIT-10 is scored by adding together the scores for all 10 items. Higher scores on this measure indicate greater alcohol misuse. It has four levels of risk for alcohol misuse: Low risk 0-7; Risky use 8-15; Harmful use 16-19; and Severe use 20-40. “At-risk” clinical status is defined as a score of 8 or higher. Items 4 to 10 refer to experiences “in the last year”. At Pre the original items and responses were used, but at the Post, these items were changed where appropriate to be “in the last 30-days.” This scale had excellent psychometrics with high internal consistency at both time points ($\alpha = .83$ Pre; $.74$ Post) and a significant test-retest correlation ($r_{\text{paired}} = .64$).

A minor methodological issue concerned the 12-month vs. the 30-day reference periods for the alcohol outcome measure at the Pre and Post periods, respectively. The response options, though, for 8 of the 10 items had 4 of the 5 choices with time frames reflecting monthly or less frequent rates of experience. In addition, our data revealed that very few responses were in the less than monthly final option on this scale, which makes the differences in response time frame for Pre and Post far less relevant.

Work Absenteeism. Developed by Chestnut Global Partners in 2010 [93], the Workplace Outcome Suite (WOS) is a psychometrically tested and validated five-scale questionnaire that has been used in over 40 EAP studies [55]. The WOS was designed to provide assessment on relevant individual differences that focus on five outcomes which are related specifically to EAP interventions and are likely to change across time if treatment is successful – and remain static if it is not. Of the five outcomes assessed by the WOS, only the absenteeism scale was used in this study. The instructions are to: “Please report for the period of the past 30 days the total number of hours your personal problems (or presenting issue)”...[item here]. It uses a fill in the blank response for the number of specific hours of absence in the past 30-days for five ways that refer to how much the employee’s ability to be at work were affected by their personal problem(s). Based on past research using the WOS [31,55,56] the following five levels were used for the severity of absence: None 0 hours; Minimal < 4 hours; Low 4-8 hours; Moderate 1-3 days (9-24 hours); High > 3 days (25-159 hours).

As other research shows the typical employee in the U.S. misses only about 3 hours per month of work due to health-related issues (see review in [31]) problem status for work absenteeism was defined as 4 or more hours of absence (4-159 hours). Based on past research using the WOS [31,55,56] we excluded cases who reported 160+ hours missed and were not actively working. Unlike the other outcome scales, the work absenteeism measure did not use a set of statements to be rated, rather it asked for specific hours of missed work to be provided in five behavioral contexts and each context is added up for total number of hours of missed work. Similar to past research [31,55,56] the total for the absenteeism scale showed a skewed pattern, with three-fourth of the total hours coming from the first item on missing a full day(s) of work and the remaining one-fourth of the total hours coming from various kinds of partial day types of absence assessed by the other four items (i.e., being late arriving to work, leaving work early, change in work location and personal communication activity while at work). The internal reliability of the WOS was not relevant to assess.

Work Productivity. Originally a 32-item version, the brief 6-item version of the Stanford Presenteeism Scale (SPS-6) is a widely used scale for assessing the impact of health problems on work productivity of employees [94,95]. According to Google scholar, since its release it has been cited in over 800 scientific papers. The SPS-6 consists of two dimensions, with one factor on completing work (items 2, 5, and 6) and a second factor on avoiding distraction while working (items 1, 3, and 4). It has response options of: (1) *Strongly disagree*; (2) *Somewhat disagree*; (3) *Uncertain*; (4) *Somewhat agree*; and (5) *Strongly agree*. The items are answered for the time period of the past month. Three of the six items are reverse scored (items 1, 3, and 4). The SPS-6 score is the sum of the three raw scores and the three reversed scores (range 6–30). A higher total scale score indicates greater work productivity despite experiencing a health issue. Based on the Consortium for Mental Healthcare [96], scale scores were grouped into five levels of work productivity, ranging from low to high: Very Low productivity 6-10; Low 11-15; Medium 16-20; High 21-25; and Very High 26-30. This scale had excellent psychometrics with high internal consistency at both time points ($\alpha = .94$ Pre; $.97$ Post) and a significant test-retest correlation ($r_{\text{paired}} = .38$). To allow us to conduct similar analyses of a change in “at-risk status” (like depression and alcohol outcomes) for this scale, we needed a way to split the distribution of SPS-6 scores into two groups of at-risk (i.e., a problem of poor productivity) or not at-risk. We defined a score on the SPS-6 at 15 or below in the very low to low range, as having a work productivity “problem”.

Note that the creators of this scale defined presenteeism as a positive aspect of work productivity [94]: “A decrease in presenteeism can hurt productivity in a way similar to an increase in absenteeism” (p. 14). However, almost all other researchers in this area define the concept of presenteeism negatively as a problem of not being psychologically present enough while working to perform properly [18-20,31,36,38-41,55,56,97,98]. For example, Cooper and Dewe [99] defined presenteeism as “lost productivity that occurs when employees come to work ill and perform below par because of that illness” (p. 522). More simply put, Hemp [100] defined presenteeism as “at work – but out of it” (p. 49). Therefore, to align the interpretation of the SPS-6 scores better with the dominant deficit-themed definition of presenteeism, we retained the original 1-5 direction of the item ratings but labelled higher scores on the SPS-6 as indicating greater work productivity while at work (i.e., better performance and greater focus while at work).

2.6. Correlation of Outcome Measures Within Cases

The four outcome measures were all correlated with each other (all $p < .001$) when tested in various samples of all available cases with data at the start of counseling. More severe depression was associated with both missing more time from work ($r = .25, n = 2,481$) and with poorer work productivity ($r = -.30, n = 2,084$). More severe alcohol misuse was related to greater work absence ($r = .23, p < .001, n = 1,485$) and slightly with lower work productivity ($r = -.10, n = 1,294$). More severe depression tended to co-occur with more severe alcohol misuse ($r = .29, n = 433$). The two work outcomes were also intercorrelated ($r = -.32, n = 25,061$). This pattern of associations in expected directions for all tests offers evidence for the convergent form of measurement validity for each outcome scale.

2.7. Data Analysis Plan

All analyses were conducted using the Statistical Package for the Social Sciences (SPSS) Version 27. Analyses with categorical variables were conducted with chi-square (χ^2) non-parametric test procedures. The tests of improvement over time (Pre to Post) in the outcome measures were conducted using a repeated measures analysis of variance procedure (RM-ANOVA). For the sub-groups of cases at-risk for depression or at-risk for alcohol misuse, we also used a more sensitive analysis developed by Jacobson and Truax [50] that tested for the extent of clinical improvement achieved based on the Reliable Change (RC) index methodology. The development of certain statistics needed for the RC index analyses were derived from data specific to our study (see Appendix A). In addition, we calculated the number of at-risk cases who demonstrated “reliable recovery” [52] which is defined as when the case experienced *both* a clinically reliable improvement (based on the size of the difference between their Pre and Post scores) and recovery on the clinical symptom measure (i.e., changed from being above the clinical cut-off score at Pre to below the threshold at Post).

2.8. Statistical Power and Effect Size

With different sample sizes for each outcome, we assessed the power to detect a particular finding as being statistically significant in the longitudinal tests [101]. The level of power to detect a small size effect in repeated measures tests at $p < .05$ chance level was very high in this study at .99 for each outcome. To allow for reasonable comparison of the findings from outcomes involving different sample sizes, we calculated the statistical effect size for most results. The partial eta squared (η_p^2) effect size statistic can range from 0 to more than 1.00, but it is usually a number closer to the zero end of the scale. These effect sizes can be interpreted as follows [102]: large size effect is $\eta_p^2 = .14$ or greater; medium size effect is $\eta_p^2 = .06$ to $.13$; small size effect is $\eta_p^2 = .01$ to $.05$; and a trivial size effect is $\eta_p^2 < .01$ (even if the finding is significant at p value less than 5% chance level). Meaningful findings in this study were defined as having both a statistically significant result and at least a small size statistical effect.

2.9 Ethical Considerations

The privacy of users was protected by having all program use and survey data deidentified before being shared with the independent consultant (first author) who conducted all statistical analyses. As this was an applied study of archival anonymized data collected from

routine use of the service, additional informed consent from individual participants beyond their initial consent agreement in terms of use of the EAP service was not required. All data was collected as part of the normal business practices and not for a separate specific research project. Project approval from a university internal review board was not required. The use and analysis of archival operational data in this manner for applied research is consistent with the published ethical guidelines of the American Psychological Association [103]. All counselors involved in the delivery of the clinical treatment services were fully licensed and trained professionals. The real-world conditions for this study are like other applied studies published in peer-review journals that have examined the effectiveness of commercial mental health support programs [44-46,56,104-107].

III. RESULTS

PART 1: Longitudinal Samples

3.1 Determination Valid Longitudinal Samples

The minimum criteria for inclusion in the longitudinal group for each outcome was having the outcome measure collected for all items on the measurement scale both at the start of the counseling and at the follow-up (Pre and Post). Employees not working were judged as irrelevant to answer questions about absence from work and productivity while working. This criterion was operationalized by excluding any employee who reported having 160 or more hours of absence in the past 30 days (we assumed a standard full-time schedule with an 8-hour daily work shift performed five days per week for the month). Excluded from the final longitudinal sample for the absenteeism outcome were 110 cases (0.4%) who were not working at Pre and 13 cases who were not working at Post (0.3%). Excluded from the final longitudinal sample for the work productivity outcome were 77 cases (0.3%) who were not working at Pre and 13 cases (0.3%) who were not working at Post. Note that cases with work productivity data who did not have work absenteeism data for the same period (i.e., at Pre or at Post) were still retained as valid cases, as the likelihood of having 160+ hours of absence - and thus not working status - for these cases was very small.

The longitudinal samples for each outcome measure are shown in Table 1. The valid longitudinal sub-groups represented from 9% to 17% of the starting sample that had data collected on the same outcome measure at the start of the case.

Table 1. Number of counseling cases with valid longitudinal data for each outcome and cases within each sample with data for the other outcome measures.

Outcome measure:	Longitudinal Sample			
	Depression	Alcohol misuse	Work absenteeism	Work productivity
Depression (PHQ-9)	487	38 (9%)	88 (2%)	93 (2%)
Alcohol misuse (AUDIT-10)	38 (8%)	440	168 (5%)	166 (5%)
Work absenteeism (WOS-5)	88 (18%)	168 (38%)	3,732	3,015 (78%)
Work productivity (SPS-6)	93 (18%)	166 (38%)	3,015 (81%)	3,854
Total valid cases	487	440	3,732	3,854
Longitudinal cases as % of total cases with data at Pre (follow-up response rate)	9% of 5,330	17% of 2,614	13% of 28,442	14% of 26,819

Note: Number of cases with primary outcome measure for sample in **bold**.

Whether an employee had completed one or more of the outcomes measures *other than* the focus for that sample (for example, cases in the depression outcome sample who also completed the work absenteeism outcome scale – see Table 1), was non-significant as a moderator of the improvement in the primary outcome from Pre to Post. Thus, if the case had also completed another outcome measure, it made no difference in the results for that primary measure.

3.2 Representativeness of Longitudinal Samples

As the cases were not randomly selected to be included in this study and the final samples included the experiences of a minority of total the service users with starting outcome information collected, we wanted to first determine if the final valid samples of cases with longitudinal data for each outcome were a reasonable representation of the much larger population of research-relevant cases at this EAP during the same period. Statistical tests were conducted comparing the Pre-only sample (i.e., those who completed a measure at baseline but not at the follow-up) with the longitudinal samples for each outcome measure. Overall, the cases of interest in the longitudinal samples for each outcome had the same general profile of the initial outcome severity level, employee demographic factors, clinical use factors, and employer context factors as other employees at this EAP (see four tables in Appendix B). Thus, the results support the representativeness of the four longitudinal samples. Few of the nine demographic, clinical, and employer context factors

differed in these tests. The few exceptions identified were for variables of referral source, clinical issue, and sex of the client. These are described below.

The percentage of cases with a formal referral by management to the EAP was very low and similar for the depression outcome longitudinal sample (3%; 16 of 487) and the Pre-only group for depression (1%; 38 of 4,843). The other three outcome longitudinal samples each had a much higher percentage of cases who were formally referred into the EAP by management at their employer than the same outcome groups that lacked the follow-up data (these differences were all medium size effects). For alcohol outcome cases, formal management referrals were 22% of the longitudinal group ($n = 97$ of 440) but only 4% of the Pre-only group ($n = 82$ of 2,174). For work absence outcome cases, formal referrals were 19% of the longitudinal group ($n = 701$ of 3,732) but less than 1% of the Pre-only group ($n = 187$ of 24,710). For work productivity outcome cases, formal referrals were 17% of the longitudinal group ($n = 644$ of 3,854) but less than 1% of the Pre-only group. Apparently, for some outcomes, being referred into the EAP by one's employer contributed to a much higher participation rate in the follow-up survey.

The depression outcome longitudinal sample did not differ from the Pre-only group on any of the demographic, clinical use or employer-related factors examined (see Appendix B, Table B1). The alcohol outcome longitudinal sample also did not differ from the Pre-only large group on any of factors examined (see Table B2), other than referral source. Other than referral source, the work absence outcome longitudinal sample only differed from the Pre-only group in one more of the factors examined (see Table B3) – type of clinical issue. Possibly influenced by additional number of the management referral cases, the absence outcome longitudinal sample did have a higher percentage of cases with presenting issues of work stress (13% > 7%), drug use (7% > 1%), and problems with alcohol use (6% > 3%) than in the Pre-only group. The work productivity outcome longitudinal sample only differed from the Pre-only group in clinical issue (see Table B4), such that the longitudinal sample had a higher percentage of cases with presenting issues of work stress (12% > 8%), drugs (5% > 2%), and alcohol (7% > 1%) than the Pre-only group.

PART 2: Changes in Outcome Measures in Longitudinal Samples

3.3 Depression Outcome Results

Depression: Change in Risk Levels. Figure 2 shows the results for how the 487 cases with a primary or secondary depression issue were distributed across the five levels of depression symptom severity at Pre and at Post. A chi-square test revealed there was a significant shift from higher to lower severity levels after counseling for many cases, $\chi^2(16,487) = 218.33, p < .001, \eta_p^2 = .28$ large effect. On an individual level, at the start of counseling 60% ($n = 292$) of employees were classified as clinically at-risk. But at Post, only 12% of cases were at the clinical risk level ($n = 56$).

Depression: Clinically At-risk Cases. Analyses were also conducted among the subsample of the 292 cases with clinical depression disorder when starting EAP counseling. The vast majority of these cases, 84% ($n = 244$) had recovered at Post to no longer be at-risk. Moreover, partial remission (a score at Post in the 5-9 range) was achieved by 45% ($n = 129$) of the at-risk cases and full remission (a score at Post in the 0-4 range) was achieved by 39% ($n = 115$) of the at-risk cases.

Depression: Reliable Change Index. Figure 3 displays the results using the RCI methodology for depression cases. It found that 89% ($n = 262$ of the 292 total cases) who started counseling at with clinical depression succeeded in having a reliable improvement (i.e., an amount of change in symptom scores from Pre to Post that was larger than chance level based on the statistical reliability of the measure). Also, 10% of the at-risk cases ($n = 28$) had no change and only 1% ($n = 2$) had a reliable deterioration in their depression symptoms (i.e., got worse). In addition, about 4 of every 5 of the at-risk depression cases (81%, $n = 236$ cases) achieved the best result of “reliable recovery” after counseling (i.e., their symptom score at Pre declined at beyond a chance amount at the Post and their Post score was below the at-risk cutoff).

Depression: Change in Average Severity Score. The RM-ANOVA test found that the level of symptom severity for the average case was 59% lower at the follow-up for all employee with longitudinal data on the PHQ-9. The same test conducted among the just the subsample who started counseling at a clinical level of depression showed an average decrease of 61% in symptom severity. Both results were significant and very large size statistical effects (see Table 2). These differences in mean score are shown in Figure 4.

Depression: Moderator Tests. Exploratory tests (RM-ANOVA) examined if the depression outcome was different for certain client demographic or clinical use context factors. The moderator results indicated only one factor – match of the clinical issue – had different outcome profiles (see Table 3). The cases who had depression as the primary reason for using counseling had more severe depression and had greater score reduction in severity of depression symptoms after use than cases with a different primary issue (anxiety, other mental health issues, personal stress, work, marriage, or other topics) yet having depression as a secondary issue. When included in the same model as clinical issue, the sex of the client was not a meaningful moderator of improvement in depression (even though women had slightly greater depression severity than men). [Note: Referral type was not tested as it did not have enough management referral cases.]

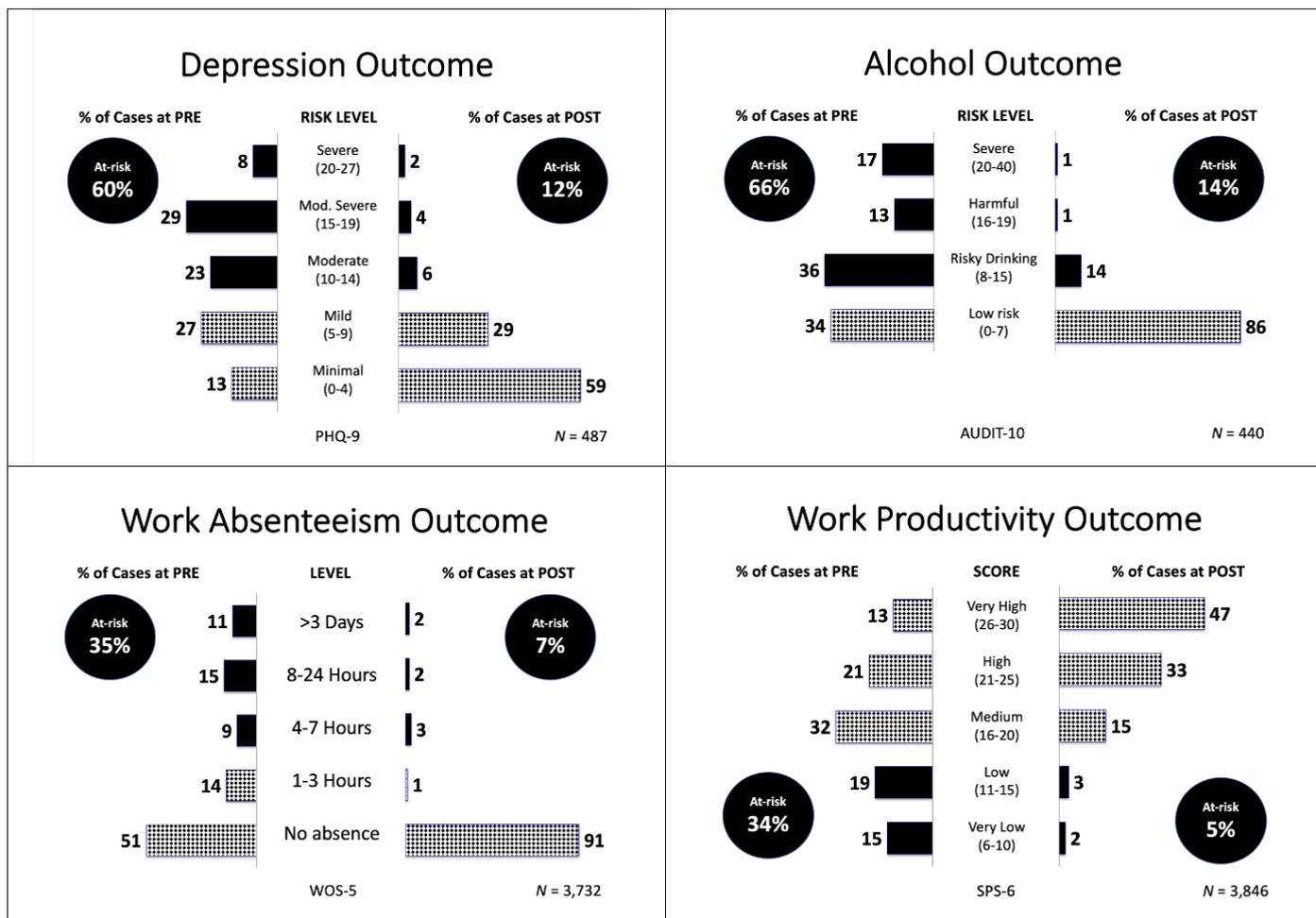


Figure 2. Percentage of cases at different levels of risk at Pre and Post in longitudinal sample: By outcome.

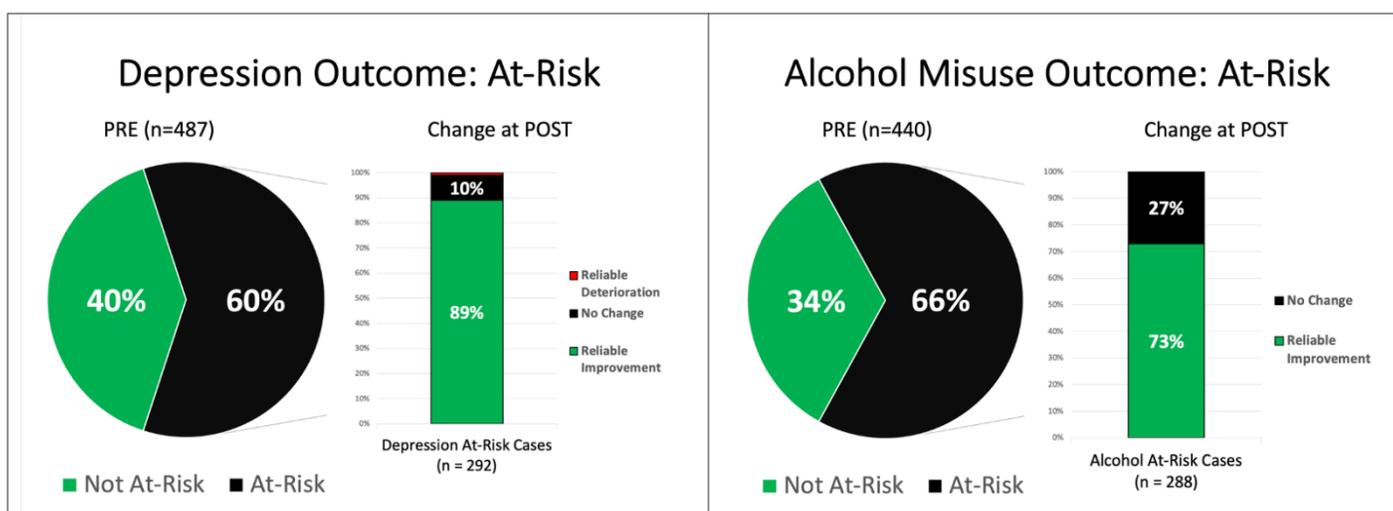


Figure 3. Reliable Change Index analysis results for cases clinically at-risk who used the EAP for a depression issue and for an alcohol issue.

Table 2. Results for longitudinal change in depression outcomes for employee users of EAP counseling.

Sample	Pre M (SD)	Post M (SD)	Change	Statistical Tests
Depression (PHQ-9)				
Full sample (<i>n</i> = 487)	11.60 (6.07)	4.73 (4.90)	-6.87 -59%	<i>Time</i> : $F(1,486) = 746.35, p < .001,$ $\eta_p^2 = .61$ very large effect
At-Risk (<i>n</i> = 292; 60%)	15.79 (3.50)	6.14 (4.31)	-9.65 -61%	<i>Time</i> : $F(1,291) = 1740.71, p < .001,$ $\eta_p^2 = .86$ very large effect
MODERATOR TESTS				
Depression issue primary (<i>n</i> = 265; 54%)	13.25 (5.47)	5.11 (4.35)	-8.14 -61%	<i>Time X Issue</i> : $F(1,485) = 32.96, p < .001,$ $\eta_p^2 = .064$ medium effect
Depression issue secondary (<i>n</i> = 222; 46%)	9.60 (6.16)	4.27 (5.47)	-5.33 -56%	<i>Issue</i> : $F(1,485) = 28.31, p < .001,$ $\eta_p^2 = .055$ small effect
Men (<i>n</i> = 223; 46%)	11.48 (0.97)	5.42 (1.18)	-6.06 -53%	<i>Time X Gender</i> : $F(1,482) = 3.07, p = .08,$ $\eta_p^2 = .006$ no effect
Women (<i>n</i> = 262; 54%)	14.07 (1.64)	5.61 (1.98)	-8.46 -60%	<i>Gender</i> : $F(1,482) = 0.03, p = .86,$ $\eta_p^2 < .001$ no effect

Factors with no effects for moderating depression outcome:

- Client sex
- Client age
- Modality of clinical use contact
- Duration of clinical treatment
- Session limit maximum
- Industry of employer
- Year of use

Note. PHQ-9 = Patient Health Questionnaire; higher scores indicate greater severity of depression experienced in the past two weeks (range 0-27). At-Risk = score 10-27. *F* tests and means for sex of client are adjusted for depression issue; with standard errors in (SE).

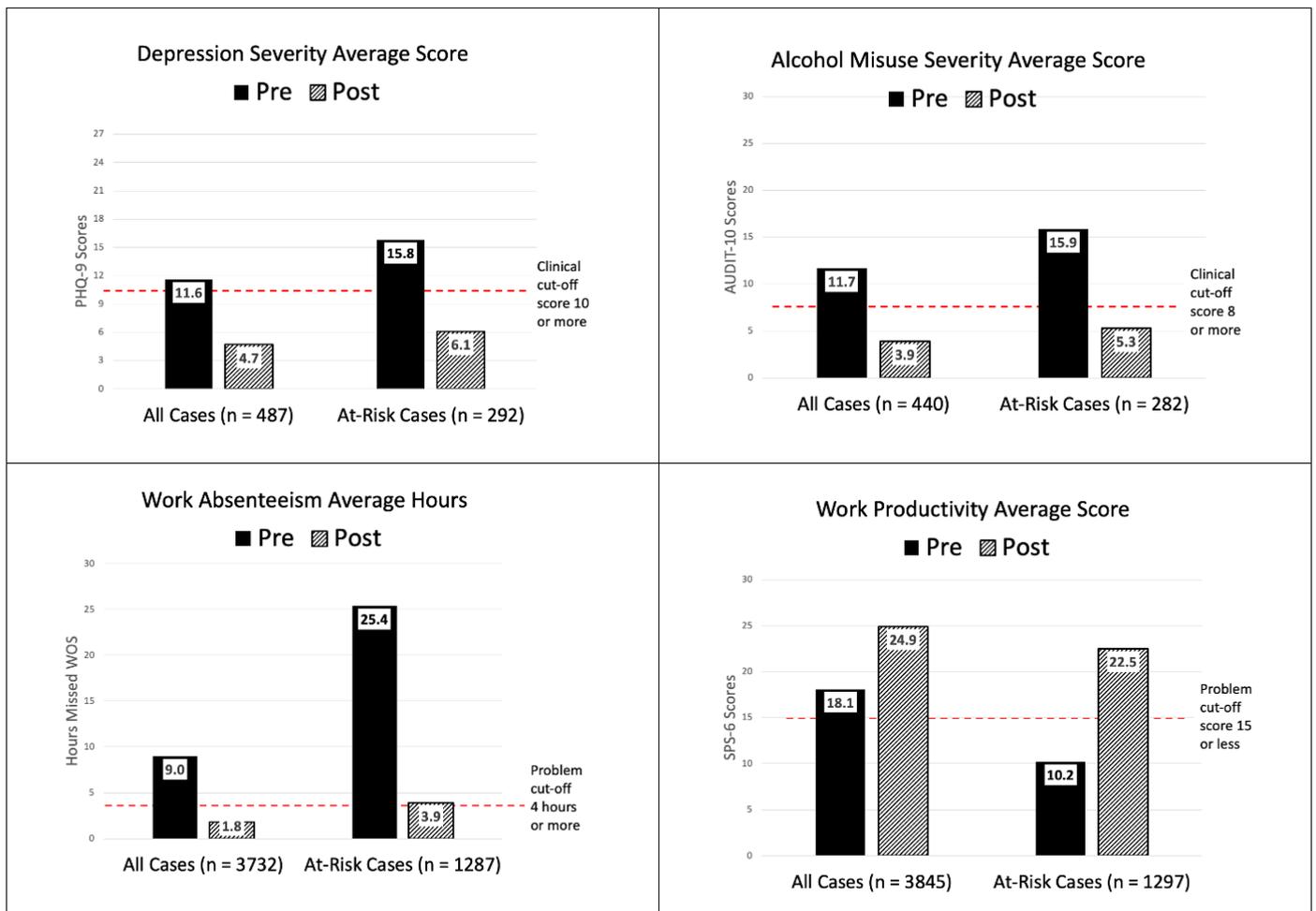


Figure 4. Mean scores at Pre and Post for full sample and at-risk subsamples: By outcome.

3.4 Alcohol Misuse Outcome Results

Alcohol: Change in Risk Levels. Figure 2 shows how the 440 cases who used the EAP for support with an alcohol issue were distributed across the four levels of alcohol misuse severity at Pre and at Post. There was a statistically significant change over time, with a shift from higher to lower severity after counseling for many cases, $\chi^2(9,440) = 81.3, p < .001$. At the start of counseling, 288 of the alcohol issue cases (66%) were initially classified as risky drinkers or at a more severe level of alcohol misuse, but at Post this number was reduced to 63 cases who were at-risk (14%).

Alcohol: Clinically At-risk Cases. Analyses were also conducted among the subsample of cases who started counseling being at-risk for alcohol misuse. For the average employee at-risk for alcohol misuse, their symptom severity was significantly lower at the follow-up. This change was a large size statistical effect (see Table 3 and Figure 4). Among individual cases, about 3 out of every 4 of the at-risk cases (78%; $n = 225$ of 288) improved to no longer be at-risk for alcohol misuse after use of EAP counseling.

Alcohol: Reliable Change Index. Figure 3 also shows the results using the RCI methodology with alcohol cases. It found that 73% of the at-risk cases had reliable improvement ($n = 210$ of 288), 27% did not have a reliable amount of change in their severity score ($n = 78$) and 0% had reliable deterioration in their alcohol misuse (see Figure 4). In addition, 175 cases had the best outcome of both a reliable reduction and improved at Post to no longer be at-risk. Thus, 61% of the at-risk alcohol cases experienced a “reliable recovery.”

Alcohol: Change in Average Severity Score. The RM-ANOVA test found that the average level of symptom severity was 67% lower at the follow-up for all employee with longitudinal data using the EAP for an alcohol issue. The same test conducted among the subsample who started counseling at a risky level of drinking or higher severity showed an average decrease of 67% in symptom severity. Both test results were significant and very large size statistical effects (see Table 3). These differences in mean score are shown in Figure 4.

Alcohol: Moderator Tests. Exploratory tests (RM-ANOVA) examined if the alcohol outcome was different for certain client demographic, clinical use, or context factors. The moderator results indicated three factors had different outcome profiles (see Table 3). For clinical issue, cases who had substance use as the primary reason for using counseling had the most severe alcohol misuse and had the greatest reduction in severity of symptoms after use than cases with a different primary issue (mental health, personal stress, etc.) with alcohol use as their secondary issue. Sex of client and referral source were also meaningful moderators in tests that controlled for alcohol as the primary issue and controlled for the other moderator factor. About twice as many men than women were had alcohol risk score data. Men also had greater severity levels of alcohol misuse and more improvement than women. For referral source, about 4 out of 5 cases who chose to use the EAP on their own choice and these cases had greater severity of alcohol misuse severity than the other 1 in 5 cases who were referred to the EAP by their employer. Self-referral cases also improved more over time than did the manager referral cases. Once the above factors were controlled for, all other factors did not have meaningful differences for the alcohol outcome.

Table 3. Results for longitudinal change in alcohol misuse outcomes for employee users of EAP counseling.

Outcome	Pre <i>M</i> (SD)	Post <i>M</i> (SD)	Change	Statistical Tests
Alcohol Misuse (AUDIT-10 scale scores)				
Full sample (<i>n</i> = 440)	11.68 (7.80)	3.93 (3.88)	-7.75 -67%	$F(1,439) = 624.15, p < .001,$ $\eta_p^2 = .58$ very large effect
At-Risk (<i>n</i> = 288; 66%)	15.90 (6.27)	5.25 (4.11)	-10.65 -67%	$F(1,287) = 857.61, p < .001,$ $\eta_p^2 = .75$ very large effect
MODERATOR TESTS				
Alcohol issue primary (<i>n</i> = 145; 33%)	15.14 (7.83)	4.99 (4.93)	-10.15 -67%	$Time \times Issue: F(1,438) = 31.80, p < .001,$ $\eta_p^2 = .068$ medium effect
Alcohol issue secondary (<i>n</i> = 295; 67%)	9.97 (7.20)	3.41 (3.11)	-6.56 -66%	$Issue: F(1,438) = 44.54, p < .001,$ $\eta_p^2 = .092$ medium effect
Men (<i>n</i> = 303; 67%)	12.51 (0.41)	4.20 (0.22)	-8.31 -66%	$Time \times Issue: F(1,436) = 7.74, p < .001,$ $\eta_p^2 = .017$ small effect
Women (<i>n</i> = 137; 31%)	9.80 (0.62)	3.32 (0.32)	-6.48 -66%	$Issue: F(1,438) = 12.50, p < .001,$ $\eta_p^2 = .028$ small effect
Referral by self (<i>n</i> = 343; 78%)	12.59 (0.39)	4.27 (0.21)	-8.32 -66%	$Time \times Referral: F(1,436) = 11.74, p < .001,$ $\eta_p^2 = .026$ small effect
Referral by manager (<i>n</i> = 97; 22%)	8.45 (0.76)	2.72 (0.40)	-5.73 -68%	$Referral: F(1,436) = 23.27, p < .001,$ $\eta_p^2 = .051$ small effect
Factors with no effects for moderating alcohol outcome:				
<ul style="list-style-type: none"> • Client age • Modality of clinical use contact • Duration of clinical treatment • Session limit maximum • Industry of employer • Year of use 				

Note. *N* = 440. AUDIT = Alcohol Use Disorders Identification Test; higher scores indicate greater severity of alcohol misuse (range 0-40). At-risk = score of 8-40. *F* tests and means for sex of client are adjusted for alcohol issue match and for referral type; with standard errors in (SE). *F* tests and means for referral are adjusted for alcohol issue match and sex of client; with standard errors in (SE).

3.5 Work Absence Outcome Results

Absence: Change in Levels of Work Absence. Figure 2 shows how the 3,732 cases who used the EAP for support with any kind of issue were distributed across the five levels of missed work at Pre and at Post. There was a statistically significant change over time, with a shift from higher to lower absence level after counseling for many cases, $\chi^2(16,3732) = 204.53, p < .001, \eta_p^2 = .048$ small. At the start of counseling, 35% of cases were classified as having an absence problem (i.e., missing more work than a typical healthy employee). At Post, only 7 % had a work absence problem.

Absence: Change in Average Hours of Absence Per Case. The RM-ANOVA test found that the average number of hours of absence in the past 30 days was 80% lower at the follow-up for all employee with longitudinal data, changing from 9 hours missed during the past 30 days at Pre to less than 2 hours during the month after counseling at Post. The same test conducted among the subsample of about 1 in 3 cases who started counseling at a problem level for absence also showed a large decrease of 85% in absenteeism level, changing from 25 hours missed at Pre to only 4 hours at Post. Both test results (using the square root transformed variable to reduce skew) were significant and large size statistical effects (see Table 4). These differences in mean score are shown in Figure 4.

Absence: Moderators Tests. Exploratory tests (RM-ANOVA) examined if the absenteeism outcome was different for certain client demographic, clinical use or context factors. The moderator results indicated two factors of referral source and issue topic had different outcome profiles in average levels of missed hours at Pre and Post (see Table 4).

There was a very wide range of average hours of absence before starting use of the counseling for cases using the EAP for different issues. The small percentage of employees with substance abuse issues involving drugs (8% of cases) or alcohol (5%) had the most absence, with both around two lost work days per month (15.4 and 14.5 hours, respectively). Cases using the EAP for depression (10%) was next with about a day and a half of missed work per month. Cases with work issues (12%) had 9.5 hours of absence per month. Cases with anxiety and other mental health issues (38%) had 8.7 hours of absence. Cases with general stress or personal life issues (10%) had 6.3 hours of absence. The lowest level of absence was found for the employees using the EAP for relationship and family issues (17%), at 4.8 hours per month. All of issue types had significant reductions after counseling, but some issues had more improvement than others (range 98% to 62% less absence). Employees with the two substance use issues had the most extreme change after counseling to average less than one hour of absence at Post.

Referral source was also meaningful moderator. For referral source, the 1 in 5 cases who were referred to the EAP by their employer had more work absence in the month before using the EAP (about 14 hours at Pre vs. 2 hours at Post) and improved more after counseling than the majority of the cases with absence data who chose to use the EAP on their own choice (about 8 hours at Pre vs. 2 hours at Post).

The combinations of certain cases for issue and referral type also indicated differences for work absenteeism levels at Pre and Post and for improvement in this outcome (see Figure 5). This interaction effect was significant [$F(4,3718) = 3.36, p = .003$] but a very small effect size ($\eta_p^2 = .004$). The cases who were referred to the EAP by their manager at work for a substance abuse issue ($n = 290$) had by far the most hours of absence during the month before counseling (17.9 hours) than all other cases. The other two groups with manager referral for mental health issues ($n = 194$) and work issues ($n = 181$) had about the same hours of missed work at Pre (10.7 and 9.6, respectively). Self-referral cases for these same three kinds of issues (substance $n = 176$; mental health $n = 1,600$; work $n = 286$) all were in same range of about 10 hours missed at Pre. The self-referral cases with personal stress ($n = 348$) and marriage/family issues ($n = 621$) were both much lower, each at around 5 hours of missed work at Pre.

Table 4. Results for longitudinal change in work absenteeism outcomes for employee users of EAP counseling.

Outcome	Pre <i>M</i> (SD)	Post <i>M</i> (SD)	Change	Statistical Tests
Work Absence (WOS)				
Full sample (<i>n</i> = 3,732)	8.98 (19.21)	1.82 (10.09)	-7.16 -80%	$F(1,3731) = 1247.22, p < .001,$ $\eta_p^2 = .25$ large effect
At-Risk (<i>n</i> = 1,287; 35%)	25.41 (25.60)	3.93 (14.84)	-10.65 -85%	$F(1,1286) = 3288.41, p < .001,$ $\eta_p^2 = .72$ very large effect
MODERATOR TESTS				
Referral by Self (<i>n</i> = 3,301; 81%)	7.89 (18.10)	2.18 (11.02)	-5.71 -72%	$Time \times Referral: F(1,3730) = 140.19, p < .001,$ $\eta_p^2 = .036$ small effect
Referral by Manager (<i>n</i> = 701; 19%)	13.70 (22.84)	1.82 (10.09)	-8.32 -66%	$Referral: F(1,3730) = 14.75, p < .001,$ $\eta_p^2 = .004$ trivial effect
Drugs (<i>n</i> = 287; 8%)	15.35 (23.85)	0.30 (2.84)	-15.05 -98%	
Alcohol (<i>n</i> = 179; 5%)	14.46 (20.33)	0.56 (2.67)	-13.90 -96%	
Depression (<i>n</i> = 392; 10%)	11.48 (25.44)	4.34 (16.66)	-7.14 -62%	$Time \times Issue: F(1,3725) = 24.97, p < .001,$ $\eta_p^2 = .039$ small effect
Work stress (<i>n</i> = 467; 12%)	9.54 (18.57)	1.38 (9.48)	-8.16 -86%	$Issue: F(6,3725) = 12.30, p < .001,$ $\eta_p^2 = .019$ small effect
Anxiety / other mental (<i>n</i> = 1403; 38%)	8.71 (18.92)	2.15 (11.13)	-6.56 -75%	
Personal stress / other (<i>n</i> = 381; 10%)	6.27 (15.85)	2.00 (9.05)	-4.27 -68%	
Marriage or family (<i>n</i> = 623; 17%)	4.76 (12.37)	0.80 (5.22)	-3.96 -83%	

Factors with no effects for moderating work absenteeism outcome:

- Client sex
- Client age
- Modality of clinical use contact
- Duration of clinical treatment
- Session limit maximum
- Industry of employer
- Year of use

Note. *N* = 3,732. WOS = Workplace Outcome Suite - Work Absence Scale; reported hours of missed work in past 30 days (range 0-159). Cases exclude if not working. At-risk = 4 or more hours of absence. *F* tests shown based on square root transformed variable of number of hours of absence (to reduce skew; means and variances for transformed data not shown).

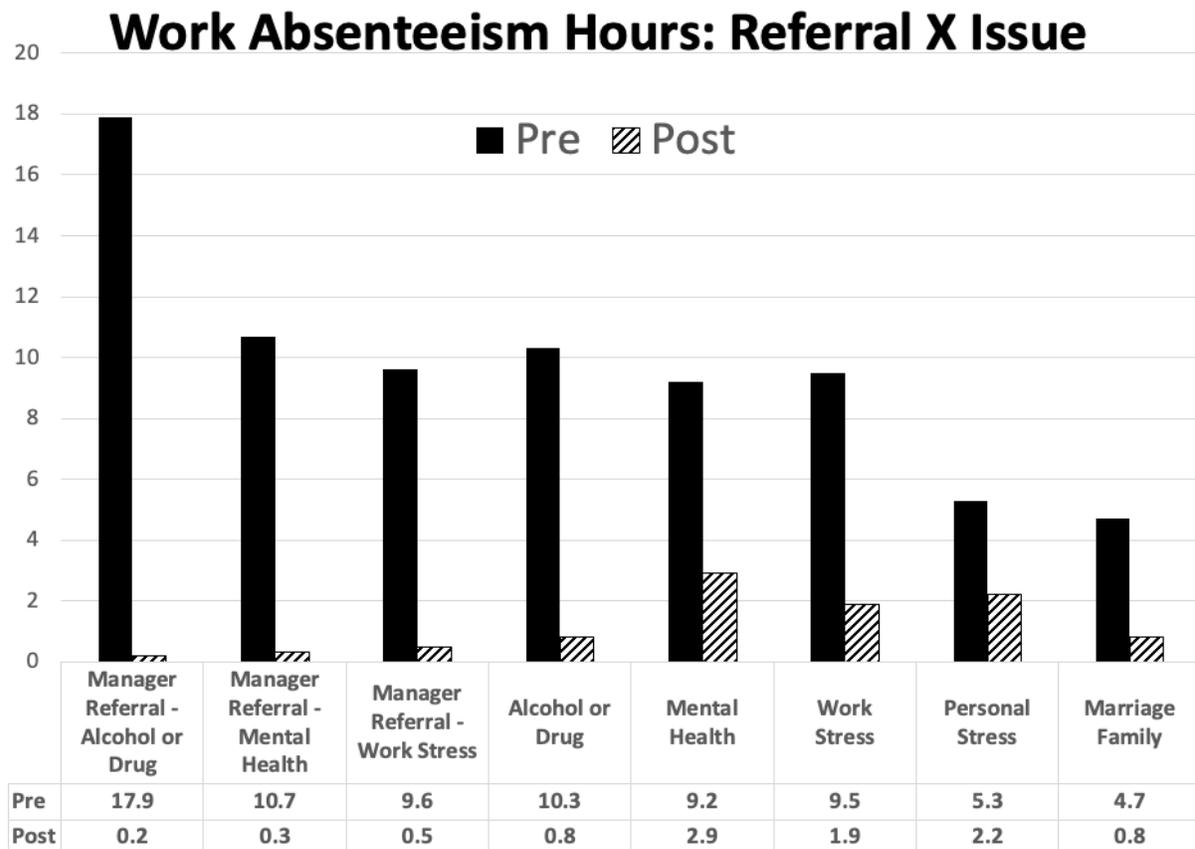


Figure 5. Mean hours of work absence at Pre and Post: By subgroups with combinations of management or self-referral into EAP and different types of issues.

3.6. Work Productivity Outcome Results

Productivity: Change in Levels of Work Productivity. Figure 2 shows how the 3,846 cases who used the EAP for support with any kind of issue were distributed across the five levels of missed work at Pre and at Post. There was a statistically significant change over time, with a shift from higher to lower absence level after counseling for many cases, $\chi^2(16,3846) = 941.60, p < .001, \eta_p^2 = .134$ medium effect. At the start of counseling, 34% cases were initially classified as having a problem with work productivity (i.e., low performance). But at Post this rate among all cases was reduced to just 5%.

Productivity: Change in Average Scores. The RM-ANOVA test found that the level of work productivity in the past 30 days for the average case was 38% higher at the follow-up for all employees with this kind of longitudinal data. The same test conducted among the just the subsample who started counseling at a problem level (about 1 in 3 cases) also showed an average increase of 38%. Both test results were significant and very large size statistical effects (see Table 5). These differences in mean score are shown in Figure 4.

Productivity: Moderator Tests. As the work productivity outcome sample shared 78% of its cases with the work absenteeism outcome sample, the results for moderator tests conducted for work productivity were very similar to the results obtained for the work absenteeism outcome sample (reported above). Exploratory tests (RM-ANOVA) examined if the work productivity outcome was different for certain client demographic, clinical use or context factors. The moderator results indicated the two factors of referral source and issue topic had different outcome profiles in average levels of productivity at Pre and Post (see Table 5). There was a very wide range of productivity before starting use of the counseling for cases using the EAP for different issues. The 4% of these cases with alcohol issues had the lowest productivity levels at Pre (the average was at the cut-off score for problem status). This was followed by cases with issues of work stress (12%), depression (11%), drugs (7%), other mental health issues (36%), and personal stress issues (10%), each of which scored in the “medium” level of work productivity at Pre. The best level of work productivity was found for cases using the

EAP for relationship and family issues (19%), with an average score in the “high” range of productivity at Pre. All of these types of issues had improvements after counseling, but some issues had more improvement than others (range 69% to 23%), generally following the same rank order as for productivity levels before counseling (i.e., those with worse productivity at the start had more opportunity to improve). For the referral source factor, the 1 in 6 cases who were referred to the EAP by their manager had slightly lower productivity during the month before using the EAP and improved significantly more after counseling than the majority of cases who chose to use the EAP on their own choice. The combinations of certain cases for issue and referral type also indicated differences for work productivity levels at Pre and Post and for improvement in this outcome (see Figure 6). This interaction effect was significant ($F[4,3835] = 3.57, p = .006$) but a very small effect size ($\eta_p^2 = .004$). Unlike with work absenteeism, these combinations of cases based on referral type and issue had more similar profiles on the work productivity outcome.

Table 5. Results for longitudinal change in work productivity outcomes for employee users of EAP counseling.

Outcome	Pre <i>M</i> (SD)	Post <i>M</i> (SD)	Change	Statistical Tests
Work Productivity (SPS-6)				
Full sample ($n = 3,845$)	18.05 (6.96)	24.87 (5.62)	6.82 38%	$F(1,3844) = 3575.94, p < .001,$ $\eta_p^2 = .482$ large effect
At-Risk ($n = 1,297; 34\%$)	10.24 (3.20)	22.53 (6.60)	12.29 38%	$F(1,1296) = 3789.09, p < .001,$ $\eta_p^2 = .745$ very large effect
MODERATOR TESTS				
Referral by Self ($n = 3,207; 83\%$)	18.31 (6.98)	24.62 (5.60)	6.31 34%	$Time \times Referral: F(1,3843) = 101.20, p < .001,$ $\eta_p^2 = .026$ small effect
Referral by Manager ($n = 638; 17\%$)	16.73 (6.71)	26.12 (5.99)	9.39 56%	$Referral: F(1,3843) = 0.40, p = .84$ $\eta_p^2 < .001$ no effect
Alcohol ($n = 170; 4\%$)	15.09 (7.54)	26.56 (4.98)	10.47 69%	
Work stress ($n = 451; 12\%$)	16.04 (5.78)	24.45 (5.57)	8.41 52%	
Depression ($n = 414; 11\%$)	16.95 (6.65)	23.65 (5.94)	6.70 40%	$Time \times Issue: F(1,3838) = 30.61, p < .001,$ $\eta_p^2 = .046$ small effect
Drugs ($n = 282; 7\%$)	17.88 (7.52)	24.97 (7.67)	7.09 40%	$Issue: F(6,3838) = 17.20, p < .001,$ $\eta_p^2 = .026$ small effect
Anxiety / other mental ($n = 1384; 36\%$)	17.96 (6.50)	25.10 (5.53)	7.14 40%	
Personal stress / other ($n = 396; 10\%$)	18.46 (7.81)	24.43 (5.48)	5.97 32%	
Marriage or family ($n = 748; 19\%$)	20.54 (6.50)	25.19 (4.77)	4.65 23%	

Factors with no effects for moderating work absenteeism outcome:

- Client sex
- Client age
- Modality of clinical use contact
- Duration of clinical treatment
- Session limit maximum
- Industry of employer
- Year of use

Note. SPS-6 = Stanford Presenteeism Scale; higher scores indicate greater focus and performance while working.

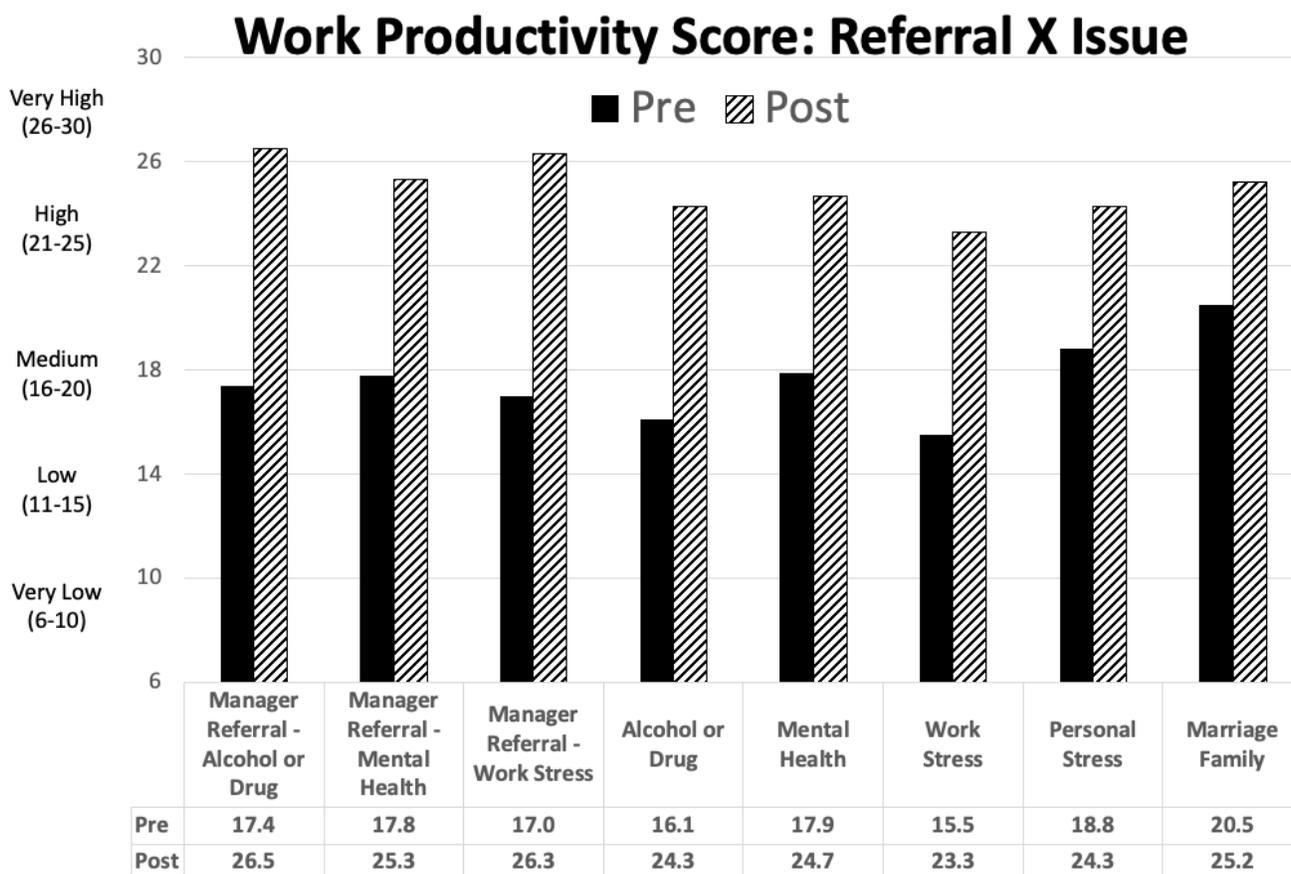


Figure 6. Average score on work productivity measure at Pre and Post: By subgroups with combinations of management or self-referral into EAP and different types of issues.

IV. DISCUSSION

Having data over a six-year calendar span from large samples of employees (over 33,000 at Pre and over 5,000 at Post) allowed us to examine the representativeness of the longitudinal test samples and the psychometric fitness of the four outcome measures. The applied naturalistic approach to the data collection process yielded groups of employee users for each outcome that had paired Pre and Post scores who fairly represented the much larger groups of cases who did not complete the follow-up survey part of the project. The outcome measures demonstrated high levels of measurement reliability and validity, with appropriate limitations for the absenteeism measure that differed in design from the other multi-item rating scales. Thus, the project provided favorable conditions to allow us to properly test for improvements in depression, alcohol misuse, work absence and work productivity after use of counseling.

4.1. Research Hypotheses Revisited

Each of the four hypotheses predicting a significant improvement in clinical or work outcomes after the use of EAP counseling were supported by the empirical findings. The average case for depression and for alcohol both had large size effects for reduction in symptom severity. More focused tests conducted only among the clinical level cases initially at-risk on depression or alcohol misuse had even greater improvement. It is important to emphasize that very few of the at-risk cases for either depression (1%) or alcohol issues (0%) had an adverse outcome of a reliable deterioration in clinical symptom severity after counseling. Thus, the counseling did no harm for almost all of the at-risk EAP users.

Both of the work outcomes also had large size statistical effects for improvement after counseling. Both work outcomes had a change from having about a third of cases starting therapy at a problem level. However, very few cases still had a problem with absence or with work productivity at the follow-up (absenteeism: 35% Pre vs. 7% Post; productivity: 34% Pre vs. 5% Post). Even though the majority of cases had zero absence at both before and after counseling, the average amount of absence per case was reduced by almost one full workday per month restored after counseling (from 9 hours at Pre to less than 2 at Post). Among the one-third of cases starting out at problem level of absence, their change was much greater, with a change from 25 to 4 hours.

One possible explanation for these positive and consistent findings is CuraLinc Healthcare's advocacy-based clinical model that drives the day to day operations. Unlike EAPs that offer participants an unverified list of mental health providers that they must contact on their own to confirm willingness to accept the referral, the advocacy-based clinical model ensured participants were connected to providers with the required clinical specialty and confirmed appointment availability. This process reduced the likelihood of failed referrals (referrals that don't lead to an appointment with a provider) which may have increased the likelihood of the study results for positive clinical and workplace outcomes, as well as the high level of client satisfaction.

4.2 Research Question on Moderators Revisited

Moderator tests indicated the improvement was generally consistent across different sub-groups of clients based on age, gender, clinical use characteristics and other study context factors. For example, for each year from 2017 to 2022 we found similar profiles of outcome levels at Pre and at Post and similar degree of improvements on each outcome at the follow-up. Similar results were also obtained for each outcome across the in-person and online video modalities of the clinical contact between counselors and clients.

Each outcome, however, did have moderator effects for clinical issue. The medium size effects found with depression and alcohol issue cases suggests the matching of primary or secondary issue (i.e., severity of the disorder) was important for testing both of these clinical outcomes. These findings support the gaps in the research literature described earlier in the Introduction concerning the value of using samples of EAP cases experiencing clinical conditions that are matched to the outcome measure for the same condition.

Other smaller size effects were also identified such that reductions in alcohol misuse severity occurred more often among cases who were self-referrals into counseling (compared to referrals from the manager at work) and also for male clients. Moderator tests found that the work outcomes had small effects for employees presenting at the EAP with different issues, with substance abuse, depression and work stress issues each being associated with more work absence and less productivity than employees with other issues (i.e., personal life stress and relationships). Note that finding that alcohol and depression issues had some of the highest levels of work loss burden before counseling supports the rationale in this project to focus on employees with alcohol or depression issues. Although accounting for a small part of the total cases, employees referred to using the EAP by their manager at work for issues involving alcohol or drugs had the most absence and also the lowest level of work productivity during the month before starting counseling. These same sub-groups also had the most improvement in these work outcomes after completing use of the EAP.

4.3 Comparisons with Other Research in EAP

The demographic and clinical use characteristics of the employees in this national study sample aligns well with other major studies of employee users of external vendors of EAP counseling services [29,31,55,56]. A strength of this study was how the representativeness of the longitudinal samples was validated by testing the user demographic characteristics, program use experiences and starting level of the outcome measures compared to the much larger group of cases with only the Pre-test data. These findings replicate those from another study that analyzed case-level data obtained from four different EAPs [see 55, Appendix C], which also found similar profiles of client age, client sex, mix of EAP issues, and initial levels of work outcomes for 3,063 cases who had longitudinal data compared with 2,681 other cases who had only Pre data. Another study compared 830 EAP cases who participated in follow-up survey with the full book of business data at the EAP and found similar profiles on all client demographic and clinical experience factors [107]. These results document the representativeness of the follow-up samples examined in some of the studies of EAP counseling effectiveness.

The present study also replicates the findings of clinical effectiveness found in several past studies of EAP counseling for cases with depression or alcohol outcomes [30,31,35-43]. More specifically, our findings for depression were compared to three other recent EAP counseling effectiveness studies [44-46] that also used the same Reliable Change Index analytic approach with the same outcome measure for depression (PHQ-9). As shown in Table 6, the results for at-risk depression cases in the present study from CuraLinc were slightly higher on all of the clinical success metrics than the findings from other providers (Lyra and Spring Health).

Our results for the work outcomes were also compared to most other studies in the literature on EAP effectiveness. The industry average is 9.7 hours absent in the past month at Pre and 4.1 hours at Post based on 6,961 counseling cases from 12 different EAPs with longitudinal data on the same five-item WOS absenteeism scale used in our study [53]. Results in the present study were 9.0 hours of work absenteeism at Pre and 1.8 hours at Post for the average EAP case. This comparison indicates that our study results were similar to the EAP industry averages for the level of work absenteeism before use and slightly better for the reduction in absence at Post.

Table 6. Comparison of present study results on depression outcome for clinically at-risk cases with three other EAP industry research studies with a similar design, outcome measure, and Reliable Change Index results.

Characteristic	Study			
	Schneider et al. [44]	Lungu et al. [45]	Bondar et al. [46]	This study
Service provider	Lyra Health	Lyra Health	Spring Health	CuraLine Healthcare
Country	United States	United States	United States	United States
Year when used	2018-2019	2018-2019	2018-2021	2017-2022
Study sample	1,256 unknown presenting issue at EAP and at least two dates of PHQ-9	225 unknown presenting issue at EAP issue and at least two dates of PHQ-9	903 unknown presenting issue at EAP and at least two dates of PHQ-9	487 with depression as a primary or secondary issue and two dates of PHQ-9
Valid sample size of EAP cases with clinical status of moderate or more severe depression (PHQ-9 = 10-27)	845 (67% of starting total)	225 (100% of starting total)	903 (100% of starting total)	292 (60% of starting total)
Intervention type	Counselors in community practice network of EAP	Blended care model of counselors in community practice network of EAP and optional dCBT	Counselors in community practice network of EAP	Blended care model of counselors in community practice network of EAP and optional dCBT
Intervention modality	Not specified (assume mostly in-person)	Online video	Not specified but a mix of in-person or online video technology	Mix of in-person (61%) or online video (39%)
Intervention intensity	Average of 9.4 sessions over 13.1 week period	Average of 5.2 sessions over 6.4 week period	Average of 5.6 sessions over 11 week period	Limit of 5.5 sessions over 5.3 week period
Design	Longitudinal with repeated measures every 4 weeks	Longitudinal with repeated measures every 4 weeks	Longitudinal pre and post	Longitudinal pre and post
Follow-up at	1 to 4 weeks post last session	63% at 1-to 4 weeks post last session and 37% before last session	Less than 4 weeks after last session	All at 4 weeks after last session
Test of representativeness of follow-up sample to larger sample with baseline measure only	Yes – similar on depression severity	Yes – similar on depression severity and also gender, age	Unknown	Yes – similar on severity of depression at Pre, age, sex, year, industry, referral, session limit, clinical duration, other
Outcome Measure	PHQ-9	PHQ-9	PHQ-9	PHQ-9
Average % reduction in severity for at-risk cases	47% lower score	48% lower score	Not shown	67% lower score
Statistical effect size for reduction in severity score	Large effect	Large effect	Large effect	Large effect
Reliable improvement beyond chance level	60% of cases	66% of cases	69% of cases	89% of cases
Recovery (no longer at-risk at post)	71% of cases	74% of cases	Not shown	84% of cases
Reliable recovery (both of above)	53% of cases	63% of cases	58% of cases	81% of cases

Note: dCBT = digital cognitive behavioral therapy for self-directed care. PHQ-9 = Patient Health Questionnaire. In each study, all tests conducted only for cases initially at-risk for clinical depression.

Our results on the work productivity outcome are more difficult to compare, as no other study in the EAP field (that we know of) has used the same Stanford Presenteeism Scale outcome measure. However, industry norms are available from over 39,000 cases with Pre and Post, from 47 different EAPs, on the Workplace Outcome Suite (WOS) single-item measure of work presenteeism [31]. These results were an average of 37% improvement in the presenteeism score after counseling, which represented a large statistical effect size.

The present study results were similar to the industry average result, as we found large 38% improvement after counseling in employee work productivity, which was also a large size statistical effect.

The present study also replicates other literature documenting the similarity in psychotherapeutic effectiveness between counseling delivered in-person when compared to counseling delivered via live video over the Internet [31,108]. This is important when the EAP industry had to adjust to offering technology-based remote options during the peak of the COVID-19 pandemic, although a return to the preference for in-person counseling was maintained during the challenges of the pandemic [109-110].

The level of individual user satisfaction with the counseling service (consistently at 95%) during the study time period is similar to the results at other EAP vendors [107]. For example, a major study [111] found an average of 94% of users satisfied with the counseling, based on results across 50 different EAP vendors in year 2011. What our results add to the literature is how satisfaction with EAP counseling was maintained during the challenges of the COVID-19 pandemic.

4.4 Limitations of the Study

There are certain limitations to this study that involve the research methodology. It was conducted on convenience samples of adults who were working for many different employers in the United States who had access to the same commercially available EAP service. Our reliance on self-report sources for the data as other records or external sources of the outcomes could have potentially provided more accurate measurements. Some studies with both sources of data, however, do show self-reports for work absence are closely aligned with company record data of employee absence [112-114]. Recent systematic reviews of the literature have also found support for the validity and reliability of employee self-reports of work presenteeism and work productivity [115,116].

Our measure of the maximum number of sessions of counseling allowed per case (from the business contract with the employer sponsor of the EAP) is not the same as the actual number of sessions of counseling used by each case. This aspect of the clinical use experience was only indirectly measured and tested.

Our use of data collected from a single group who used the counseling with repeated measures of outcomes is an example of a “pre-experimental” type of research design with no comparison group. The other two stronger kinds of designs are quasi-experimental (with a comparison group involved) and the experimental randomized control trial (RCT). We also did not measure the possible use of other health-enhancing treatments during the same period as when the EAP was used. Therefore, the causal mechanism of how use of the EAP was specifically related to the changes reported on the clinical and work outcome measures needs further scrutiny to determine what are the most impactful aspects of the clinical experience and supportive resources. However, to be fair, in the entire 90-year history of the EAP field, we are aware of no study that has used a true randomized control trial (RCT) design and only a handful of studies have used a quasi-experimental longitudinal study design with a matched comparison group [40-42] or other comparison groups. In McLeod’s 2008 review [62], five studies involving comparison groups all found that the EAP treatment group had superior improvement over time on clinical outcomes compared to other employees who did not use counseling. The two recent quasi-experimental small sample EAP studies with matched comparison groups of non-users, also both found superior improvements in most clinical and work outcomes for the EAP user groups. Thus, even though our study lacked a comparison sample, other past studies with comparison groups have determined that the EAP counseling users had superior outcomes compared to non-users of EAP.

V. CONCLUSION

This applied evaluation study provides evidence that brief counseling is associated with large size improvements for employees in both behavioral health and work domains. This study also shows that applied naturalistic research can be done in business settings that provide counseling. Other providers of EAP services are encouraged to also adopt the use of research-validated measurement scales for clinical and work outcomes and to conduct the testing necessary to ensure that longitudinal samples with outcomes are representative of the larger overall case mix at the EAP. Like some past research [31,35,46,55,61,70-72,107], future analyses could be done to estimate the potential cost-savings to employer sponsors of EAP services in areas of health care treatment costs for depression and alcohol and also for reduced absenteeism and restored work productivity.

APPENDIX A

A.1. Reliable Change Index for Depression Outcome

For the 292 cases with moderate or more severe symptoms of depression, the Reliable Change (RC) index was calculated. We used the study-specific standard deviation (SD) of the PHQ-9 at Pre of 3.50. However, the test-retest correlation of PHQ-9 scores within our at-risk sample ($r = .50$) was biased downward by the counseling intervention experience that occurred between Pre and Post. Following Hageman and Arrindel [51]), the Cronbach alpha coefficient for the depression scale (full study sample at Pre, $n = 5,331$, $\alpha = .79$) was used instead of the test-retest r for the reliability level for the PHQ-9 scale. Note that α in our study was similar to the test-retest correlations found in prior studies with short-term re-assessment of PHQ-9 scores in clinical populations ($r = .81$ to $.96$, [83,84]) and in normal adult populations ($r = .86$, [86]) that did not have an intervention between Time 1 and 2. Using these inputs, the RC index for depression was calculated as follows for the Standard Error (SE) and the difference in SE (SE_{diff}):

SD of PHQ-9 scores at Pre = 3.50
Reliability of PHQ-9 scores at Pre = .79
 $SE = 3.50 \times (\text{square root of } 1.00 - 0.79)$
 $SE = 3.50 \times (\text{square root of } 0.21)$
 $SE = 3.50 \times 0.458$
 $SE = 1.60$
 $SE \text{ squared} = 1.60 \times 1.60 = 2.57$
 $SE_{diff} = \text{square root of } (2 \times SE \text{ squared})$
 $SE_{diff} = \text{square root of } (2 \times 2.57)$
 $SE_{diff} = \text{square root of } 5.14$
 $SE_{diff} = 2.27$

RC index score for each case in the group at-risk for depression was then calculated in three steps: 1) taking the difference between the Pre and Post scores on the PHQ-9 measure; 2) dividing that difference by the RC index of 2.27; and 3) comparing that result to the probability chance-level cutoff of 1.96. If a case had RC effect score greater than 1.96, that client had a reliable improvement (or a reliable deterioration for a negative score > -1.96). For example, a case had a Pre score of 20 and a Post score of 12, the difference was 8. When this difference was divided by RC index of 2.27 the result is 3.52. This final number is greater than the 1.96 probability cutoff and thus this case is declared as having a reliable improvement.

B.2. Reliable Change Index for Alcohol Misuse Outcome

For the 288 cases with risky drinking or more severe alcohol misuse, the Reliable Change index was calculated. We used the study-specific at-risk subsample's standard deviation of the AUDIT-10 at Pre of 6.27. However, the test-retest correlation of PHQ-9 scores within the at-risk sample ($r = .35$) in our study was biased downward by the counseling intervention experienced in between the Pre and Post. The Cronbach alpha reliability coefficient for the AUDIT-10 scale for the Pre-only sample ($n = 2,174$; $\alpha = .87$) was used instead as the reliability level for the AUDIT-10 to calculate the RC index. Note that this reliability coefficient in our study was very similar to test-retest correlations found in other studies ($r = .86$ [90]). Using these inputs, the RC index for alcohol was calculated in a similar manner as for depression, with data inputs as follows:

SD of AUDIT-10 scores at Pre for at-risk cases = 6.27
Reliability of AUDIT-10 at Pre for at-risk cases = 0.87
 $SE = 6.27 \times (\text{square root of } 1.00 - 0.87)$
 $SE = 6.27 \times (\text{square root of } 0.13)$
 $SE = 6.27 \times 0.360$
 $SE = 2.26$
 $SE \text{ squared} = 2.26 \times 2.26$
 $SE \text{ squared} = 5.09$
 $SE_{diff} = \text{square root of } 2 \times SE \text{ squared}$
 $SE_{diff} = \text{square root of } 2 \times 5.09$
 $SE_{diff} = \text{square root of } 10.19$
 $SE_{diff} = 3.19$

RC index score for each case in the group at-risk for alcohol misuse was then calculated in three steps: 1) taking the difference between the Pre and Post scores on the AUDIT-10 measure; 2) dividing that difference by the RC index of 3.19; and 3) comparing that result to the probability chance-level cutoff of 1.96.

APPENDIX B

Table B1. Depression outcome: Tests of representativeness of longitudinal sample.

	Group		Statistical Test	
	Pre only	Longitudinal		
<i>n</i> cases	4,844	487		
Depression outcome (PHQ-9 at Pre)			$F(1,5330) = 6.59, p = .01,$	
<i>M</i> (SD)	10.9 (5.3)	11.6 (6.1)	$\eta_p^2 < .001$ no effect	
Clinical issue as primary focus during treatment				
Depression (outcome)	46%	56%	$\chi^2(8,5331) = 80.34, p < .001,$ $\eta_p^2 = .015$ small effect	
Personal stress	12%	11%		
Other MH	12%	10%		
Anxiety	18%	6%		
Work stress	3%	5%		
Marital	4%	4%		
Family	3%	3%		
Alcohol	2%	1%		
Drugs	1%	3%		
Referral source into EAP				
Self	99%	97%	$\chi^2(1,5331) = 127.61, p < .001,$ $\eta_p^2 = .005$ no effect	
Management	1%	3%		
Age of client				
Years <i>M</i> (SD)	38.8 (12.5)	35.7 (10.9)	$F(1,2486) = 27.45, p = .03,$ $\eta_p^2 = .005$ no effect	
	<i>n</i> = 4771	<i>n</i> = 479		
Sex of client				
Male	37%	46%	$\chi^2(1,5306) = 16.51, p < .001,$ $\eta_p^2 = .002$ no effect	
Female	63%	54%		
Clinical modality				
In-person	68%	61%	$\chi^2(2,5331) = 12.62, p = .002,$ $\eta_p^2 = .002$ no effect	
Online video	32%	39%		
Other technology	<1%	-		
Year of data				
2017	5%	6%	$\chi^2(5,5331) = 287.88, p < .001,$ $\eta_p^2 = .05$ small effect	
2018	9%	6%		
2019	7%	18%		
2020	12%	34%		
2021	52%	27%		
2022 Jan-June.	15%	9%		
Industry of employer				
Manufacturing	18%	23%	$\chi^2(8,5331) = 24.50, p < .001,$ $\eta_p^2 = .005$ no effect	
Healthcare	25%	21%		
Education	13%	13%		
Retail/Restaurant	5%	12%		
Other	7%	10%		
Gov./Public Service	8%	6%		
Financial/Insurance	9%	6%		
Blue collar	6%	5%		
Technology	7%	3%		
Maximum sessions allowed per case in employer contract				
3 or 4 sessions	12%	9%		$F(1,5330) = 2.54, p = .11,$ $\eta_p^2 < .001$ no effect
5 sessions	43%	47%		
6 sessions	26%	28%		
7 or more sessions	19%	15%		
<i>M</i> (SD)	5.8 (1.8)	5.7 (1.5)		

Note. PHQ = Patient Health Questionnaire; higher scores indicate greater severity of depression experienced in the past two weeks (range 0-27).

Table B2. Alcohol misuse outcome: Tests of representativeness of longitudinal sample.

	Group		Statistical Test
	Pre only <i>n</i> cases	Longitudinal	
Alcohol misuse outcome (AUDIT-10 at Pre)	2,174	440	
<i>M</i> (SD)	12.0 (8.9)	11.7 (7.8)	$F(1,2612) = 0.36, p = .55,$ $\eta_p^2 < .001$ no effect
Clinical issue as primary focus during treatment			
Alcohol (outcome)	37%	35%	$\chi^2(8,2614) = 27.10, p < .001,$ $\eta_p^2 < .001$ no effect
Depression	13%	15%	
Other MH	10%	13%	
Personal stress	12%	11%	
Anxiety	13%	11%	
Work stress	3%	6%	
Marital	7%	5%	
Drugs	2%	4%	
Family	2%	1%	
Referral source into EAP			
Self	96%	78%	$\chi^2(1,2614) = 191.56, p < .001,$ $\eta_p^2 = .073$ medium effect
Management	4%	22%	
Age of client			
Years <i>M</i> (SD)	40.5 (11.5)	42.2 (12.8)	$F(1,2581) = 7.70, p = .01,$ $\eta_p^2 = .003$ no effect
<i>n</i>	2,150	433	
Sex of client			
Male	61%	69%	$\chi^2(1,2604) = 10.53, p < .001,$ $\eta_p^2 = .004$ no effect
Female	39%	31%	
Clinical modality			
In-person	64%	62%	$\chi^2(2,2614) = 2.15, p = .34,$ $\eta_p^2 < .001$ no effect
Online video	35%	37%	
Other technology	1%	<1%	
Year of data			
2017	5%	6%	$\chi^2(5,5331) = 287.88, p < .001,$ $\eta_p^2 = .05$ small effect
2018	9%	6%	
2019	7%	18%	
2020	12%	34%	
2021	52%	27%	
2022 Jan-June.	15%	9%	
Industry of employer			
Manufacturing	30%	31%	$\chi^2(8,2614) = 35.27, p < .001,$ $\eta_p^2 = .003$ no effect
Blue collar	8%	13%	
Healthcare	14%	17%	
Education	9%	10%	
Gov./Public Service	9%	4%	
Financial/Insurance	8%	4%	
Retail/Restaurant	11%	11%	
Technology	4%	4%	
Other	7%	6%	
Maximum sessions allowed per case in employer contract			
3 or 4 sessions	15%	8%	$F(1,2613) = 1.05, p = .31,$ $\eta_p^2 < .001$ no effect
5 sessions	47%	50%	
6 sessions	26%	32%	
7 or more sessions	13%	10%	
<i>M</i> (SD)	5.4 (1.6)	5.5 (1.3)	

Note. AUDIT = Alcohol Use Disorders Identification Test; higher scores indicating greater severity of alcohol misuse.

Table B3. Work absenteeism outcome: Tests of representativeness of longitudinal sample.

	Group		Statistical Test
	Pre only <i>n</i> cases	Longitudinal	
Work absenteeism hours outcome (WOS at Pre)	24,697	3,733	
Hours <i>M</i> (SD)	5.7 (14.1)	9.0 (19.2)	$F(1,28242) = 1218.62, p < .001,$ $\eta_p^2 < .008$ no effect
SQRT <i>M</i> (SD)	1.2 (2.1)	1.8 (2.4)	
Clinical issue as primary focus during treatment			
Other MH	18%	19%	
Anxiety	18%	18%	
Marital	16%	13%	
Work stress (outcome)	7%	13%	$\chi^2(8,28425) = 1246.97, p < .001,$ $\eta_p^2 = .044$ small effect
Depression	14%	11%	
Personal stress	19%	10%	
Drugs	1%	7%	
Alcohol	3%	6%	
Family	5%	4%	
Referral source into EAP			
Self	99%	81%	$\chi^2(1,28430) = 3480.50, p < .001,$ $\eta_p^2 = .123$ medium effect
Management	1%	19%	
Age of client			
Years <i>M</i> (SD)	39.6 (11.4)	40.2 (11.6)	$F(1,27367) = 9.38, p = .002,$ $\eta_p^2 < .001$ no effect
	<i>n</i> = 23,836	<i>n</i> = 3,532	
Sex of client			
Male	37%	45%	$\chi^2(1,28321) = 86.81, p < .001,$ $\eta_p^2 = .003$ no effect
Female	63%	55%	
Clinical modality			
In-person	73%	78%	$\chi^2(2,28430) = 69.90, p < .001,$ $\eta_p^2 = .003$ no effect
Telephone	24%	18%	
Other technology	3%	4%	
Year of data			
2017	6%	10%	
2018	16%	21%	
2019	19%	19%	$\chi^2(5,28430) = 172.99, p < .001,$ $\eta_p^2 = .006$ no effect
2020	26%	25%	
2021	26%	21%	
2022 Jan-June.	8%	6%	
Industry of employer			
Manufacturing	20%	23%	
Healthcare	21%	17%	
Education	13%	11%	
Gov./Public Service	11%	11%	$\chi^2(8,28430) = 106.23, p < .001,$ $\eta_p^2 = .004$ no effect
Financial/Insurance	10%	8%	
Technology	6%	8%	
Retail/Restaurant	10%	9%	
Blue collar	5%	7%	
Other	5%	6%	
Maximum sessions allowed per case in employer contract			
3 or 4 sessions	13%	11%	
5 sessions	45%	49%	$F(1,28429) = 13.63, p < .001,$ $\eta_p^2 < .001$ no effect
6 sessions	27%	29%	
7 or more sessions	15%	11%	
<i>M</i> (SD)	5.6 (1.6)	5.5 (1.4)	

Note. WOS = Workplace Outcome Suite; higher scores indicate more absence (range 0-159 hours). Absence was transformed with a square root procedure (SQRT) for statistical testing; but the reported hours of absence are shown in this table.

Table B4. Work productivity outcome: Tests of representativeness of longitudinal sample.

	Group		Statistical Test
	Pre only <i>n</i> cases	Longitudinal	
Work productivity outcome (SPS-6 at Pre)	22,941	3,846	$F(1,26786) = 24.96, p < .001,$ $\eta_p^2 < .001$ no effect
<i>M</i> (SD)	18.7 (7.5)	18.1 (7.0)	
Clinical issue as primary focus during treatment			
Other MH	18%	18%	$\chi^2(8,26783) = 1117.38, p < .001,$ $\eta_p^2 = .042$ small effect
Anxiety	18%	17%	
Marital	16%	15%	
Depression	13%	12%	
Work stress (outcome)	8%	12%	
Personal stress	19%	10%	
Alcohol	2%	5%	
Drugs	<1%	7%	
Family	5%	5%	
Referral source into EAP			
Self	99%	83%	$\chi^2(1,26787) = 2893.13, p < .001,$ $\eta_p^2 = .108$ medium effect
Management	1%	17%	
Age of client			
Years <i>M</i> (SD)	39.6 (11.5)	40.2 (11.4)	$F(1,25320) = 9.48, p = .002,$ $\eta_p^2 < .001$ no effect
<i>n</i>	21,765	3,556	
Sex of client			
Male	37%	46%	$\chi^2(1,26693) = 117.92, p < .001,$ $\eta_p^2 = .004$ no effect
Female	63%	54%	
Clinical modality			
In-person	74%	78%	$\chi^2(2,26787) = 30.00, p = .34,$ $\eta_p^2 < .001$ no effect
Telephone	23%	19%	
Other technology	3%	3%	
Year of data			
2017	6%	8%	$\chi^2(5,26787) = 59.51, p < .001,$ $\eta_p^2 = .002$ no effect
2018	17%	19%	
2019	20%	20%	
2020	27%	26%	
2021	24%	21%	
2022 Jan-June.	7%	6%	
Industry of employer			
Manufacturing	19%	23%	$\chi^2(8,26787) = 140.39, p < .001,$ $\eta_p^2 = .005$ no effect
Blue collar	5%	7%	
Healthcare	21%	17%	
Education	13%	10%	
Gov./Public Service	11%	11%	
Financial/Insurance	10%	8%	
Retail/Restaurant	9%	10%	
Technology	7%	9%	
Other	5%	6%	
Maximum sessions allowed per case in employer contract			
3 or 4 sessions	13%	10%	$F(1,26786) = 6.12, p = .01,$ $\eta_p^2 < .001$ no effect
5 sessions	45%	49%	
6 sessions	27%	29%	
7 or more sessions	15%	12%	
<i>M</i> (SD)	5.6 (1.7)	5.5 (1.4)	

Note. SPS = Stanford Presenteeism Scale; higher scores indicate greater work productivity and focus while at work in past month (range 6-30).

ACKNOWLEDGMENT

We acknowledge support of the many counselors and other staff at CuraLinc Healthcare who provided the clinical services and collected the data analyzed in this project. We are grateful to the thousands of employees who voluntarily used the counseling services for making the study possible.

DECLARATIONS

Funding: The research was financed by the authors' own resources. No external research grant funding was involved.

Author Contributions: MA performed the statistical analyses of the aggregated dataset, conducted the literature review and drafted the manuscript. DP and SF developed the study design, selected the measures involved, coordinated the data collection and led preparation of annual reports of preliminary results. All authors discussed the results and contributed to the final manuscript.

Conflict of interest/Competing interests: MA is an independent research scholar and consultant who received financial support from CuraLinc Healthcare for preparing this research manuscript. MA has also occasionally worked on other projects for this company. DP and SF work for CuraLinc Healthcare company.

Institutional Review Board Statement: No formal ethical approval of the study was required due to the retrospective archival naturalistic design of the study. All employees who used the counseling and completed the outcome measures participated voluntarily and had their personal identity protected as all unique identifiers were removed from the data prior to analysis. All counselors involved in the delivery of the clinical treatment services were fully licensed and trained professionals. All aspects of this evaluation project and preparation of the manuscript followed the ethical guidelines of the American Psychological Association (2017).

Informed Consent Statement: All data was collected as part of the normal business practices and not for a separate specific research project. Consent for participation in a research study and use of data for publication of study results was therefore not necessary.

REFERENCES

- [1] Attridge, M. (2019). A global perspective on promoting workplace mental health and the role of employee assistance programs. *American Journal of Health Promotion*, 33(4), 622-29. <https://doi.org/10.1177/0890117119838101c>
- [2] Karg, R.S., Bose, J., Batts, K.R., Forman-Hoffman, V.L., Liao, D., Hirsch, E., et al. (2014). Past year mental disorders among adults in the United States, Results from the 2008–2012 Mental Health Surveillance Study. *CBHSQ Data Review*. Center for Behavioral Health Statistics and Quality, SAMHSA. <https://europaemc.org/article/MED/27748100>
- [3] Substance Abuse and Mental Health Services Administration. (2019). *Key substance use and mental health indicators in the United States, results from the 2018 National Survey on Drug Use and Health*. <https://www.samhsa.gov/data/report/2018-nsduh-annual-national-report>
- [4] Panchal, N., Kamal, R., Orgera, K., Cox, C., Garfield, R., Hamel, L., et al. (2021). *The implications of COVID-19 for mental health and substance use* [White paper]. Kaiser Family Foundation. https://pameladwilson.com/wp-content/uploads/4_5-2021-The-Implications-of-COVID-19-for-Mental-Health-and-Substance-Use_-_KFF-1.pdf
- [5] Abdalla, S.M., Ettman, C.K., Cohen, G.H., Galea, S. (2021). Mental health consequences of COVID-19: A nationally representative cross-sectional study of pandemic-related stressors and anxiety disorders in the USA. *BMJ Open*, 11, e044125. doi:10.1136/bmjopen-2020-044125
- [6] Xiong, J., Lipsitz, O., Nasri, F., Lui, L.M., Gill, H., Phan, L., et al. (2020). Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *Journal of Affective Disorders*, 277(1), 55-64. <https://doi.org/10.1016/j.jad.2020.08.001>
- [7] Kessler, R. C., de Jonge, P., Shahly, V., van Loo, H. M., Wang, P. S.-E., & Wilcox, M. A. (2014). Epidemiology of depression. In I. H. Gotlib & C. L. Hammen (Eds.), *Handbook of Depression* (pp. 7–24). Guilford Press.
- [8] Hawton, K., Icc, C., Haw, C., & Saunders, K. (2013). Risk factors for suicide in individuals with depression: A systematic review. *Journal of Affective Disorders*, 147(1–3), 17–28. <https://doi.org/10.1016/j.jad.2013.01.004>
- [9] Proudman, D., Greenberg, P., & Nellesen, D. (2021). The growing burden of major depressive disorders (MDD): Implications for researchers and policy makers. *Pharmacoeconomics*, 39(6), 619-625. <https://doi.org/10.1007/s40273-021-01040-7>
- [10] Greenberg, P.E., Fournier, A.A., Sisitsky, T., Simes, M., Berman, R., Koenigsberg, S.H., & Kessler, R.C. (2021). The economic burden of adults with major depressive disorder in the United States (2010 and 2018). *Pharmacoeconomics*, 39(6), 653-665. <https://doi.org/10.1007/s40273-021-01019-4>
- [11] Herrman, H., Kieling, C., McGorry, P., Horton, R., Sargent, J., & Patel, V. (2019). Reducing the global burden of depression: A Lancet–World Psychiatric Association Commission. *The Lancet*, 393(10189), e42-e43. [https://doi.org/10.1016/S0140-6736\(18\)32408-5](https://doi.org/10.1016/S0140-6736(18)32408-5)
- [12] U.S. Department of Health and Human Services (HHS), Office of the Surgeon General. (2016). *Facing addiction in America: The Surgeon General's report on alcohol, drugs and health*. <https://store.samhsa.gov/sites/default/files/d7/priv/surgeon-generals-report.pdf>
- [13] Stahre, M., Roeber, J., Kanny, D., Brewer, R.D., & Zhang, X. (2014). Contribution of excessive alcohol consumption to deaths and years of potential life lost in the United States. *Preventing Chronic Disorders*, 11(E109). <https://dx.doi.org/10.5888/pcd11.130293>
- [14] Mekonen, T., Chan, G.C., Connor, J.P., Hides, L., & Leung, J. (2021). Estimating the global treatment rates for depression: A systematic review and meta-analysis. *Journal of Affective Disorders*, 295(1), 1234-1442. <https://doi.org/10.1016/j.jad.2021.09.038>

- [15] O'Donnell, M.P., Schultz, A.B., & Yes, L. (2015). The portion of health care costs associated with lifestyle-related modifiable health risks based on a sample of 223,461 employees in seven industries: The UM-HMRC Study. *Journal of Occupational and Environmental Medicine*, 57, 1284-1290. doi:10.1097/JOM.0000000000000600
- [16] Goetzel, R.Z., Henke, R.M., Head, M.A., Benevent, R., & Rhee, K. (2020). Ten modifiable health risk factors and employees' medical costs: An update. *American Journal of Health Promotion* 34(5), 490-499. <https://doi.org/10.1177/0890117120917850>
- [17] Frey, J.J., Osteen, P.J., Berglund, P.A., Jinnett, K., & Ko, J. (2015). Predicting the impact of chronic health conditions on workplace productivity and accidents: Results from two US Department of Energy national laboratories. *Journal of Occupational and Environmental Medicine*, 57, 436-444. doi:10.1097/JOM.0000000000000383
- [18] Allen, D., Hines, E.W., Pazdernik, V., Konecny, L.T., & Breitenbach, E. (2018). Four-year review of presenteeism data among employees of a large United States health care system: A retrospective prevalence study. *Human Resources for Health*, 16(1), 1-10. <https://doi.org/10.1186/s12960-018-0321-9>
- [19] Thørrisen, M.M., Bonsaksen, T., Hashemi, N., Kjekens, I., Van Mechelen, W., & Aas, R.W. (2019). Association between alcohol consumption and impaired work performance (presenteeism): A systematic review. *BMJ Open*, 9(7):e029184. <https://doi.org/10.1136/bmjopen-2019-029184>
- [20] Burton, W.N., Chen, C.Y., Conti, D.J., Schultz, A.B., & Edington, D.W. (2006). The association between health risk change and presenteeism change. *Journal of Occupational and Environmental Medicine*, 48(3), 252-263. <https://doi.org/10.1097/01.jom.0000201563.18108.af>
- [21] Mitchell, R.J., Ozminowski, R.J., & Serxner, S. (2013). Improving employee productivity through improved health. *Journal of Occupational and Environmental Medicine*, 55(10), 1142-1148. <https://www.jstor.org/stable/48500317>
- [22] Attridge, M., & Dickens, S.P. (2021). Onsite screening and enhanced EAP counseling improves overall health, depression, and work outcomes: Four-wave longitudinal pilot study at a community health center in Vermont. *Journal of Workplace Behavioral Health*, 36(4), 278-308. <https://doi.org/10.1080/15555240.2021.1971537>
- [23] Goetzel, R.Z., Roemer, E.C., Hologue, C., Fallin, M.D., McCleary, K., Eaton, W., Agnew, J., Azocar, F., Ballard, D., Bartlett, J., & Braga, M. (2018). Mental health in the workplace: A call to action proceedings from the mental health in the workplace: Public health summit. *Journal of Occupational and Environmental Medicine*, 60(4), 322-330. doi:10.1097/JOM.000000000000127
- [24] Aarons-Mele, M. (2018, Nov. 1). We need to talk more about mental health at work. *Harvard Business Review*. <https://hbr.org/2018/11/we-need-to-talk-more-about-mental-health-at-work>
- [25] Pfeffer, J., & Williams, L. (2020, Dec. 8). Mental health in the workplace: The coming revolution. *McKinsey Quarterly*. www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/mental-health-in-the-workplace-the-coming-revolution#
- [26] Society for Human Resources Management, SHRM Foundation, & Ostuka Pharmaceutical Co. (2022). *Mental health in America: A 2022 workplace report*. <https://www.workplacementhealth.shrm.org/wp-content/uploads/2022/04/Mental-Health-in-America-A-2022-Workplace-Report.pdf>
- [27] Employee Assistance Professionals Association (2015). *Orientation to Employee Assistance Programs: For Mental Health Professionals in the U.S.A.* [Presentation]. <http://hdl.handle.net/10713/15775>
- [28] Bureau of Labor Statistics, US Department of Labor, United States Government (2021). *National Compensation Survey: Employee Benefits in the United States, March 2021*. <https://www.bls.gov/ncs/ebs/benefits/2021/employee-benefits-in-the-united-states-march-2021.pdf>
- [29] Prottas, D.J., Diamante, T., & Sandys, J. (2011). The U.S. domestic workforce use of employee assistance services: An analysis of ten years of calls. *Journal of Workplace Behavioral Health*, 26(4), 296-211. <https://doi.org/10.1080/15555240.2011.618431>
- [30] Lam, R.W., Wolinsky, D., Kinsella, C., Woo, C., Cayley, P.M., & Walker, A.B. (2012). The prevalence and impact of depression in self-referred clients attending an employee assistance program. *Journal of Occupational and Environmental Medicine*, 54(11), 1395-1399. doi:10.1097/JOM.0b013e3182611a69
- [31] Attridge M. (2022). *Workplace Outcome Suite (WOS) Annual Report 2021: EAP counseling use and outcomes, COVID-19 pandemic impact, and best practices in outcome data collection* [White paper]. LifeWorks. <http://hdl.handle.net/10713/18701>
- [32] Roman, P.M., & Blum, T.C. (2004). Employee assistance programs and other workplace preventive strategies. In Galanter, M., & Kleber, H.D. (Eds.), *The textbook of substance abuse treatment, 3rd Edn.* (pp. 423-435). American Psychiatric Association Press.
- [33] Jacobson, J.M., & Sacco, P. (2012). Employee assistance program services for alcohol and other drug problems: Implications for increased identification and engagement in treatment. *American Journal of Addictions*, 21, 468-475. <https://dx.doi.org/10.1111/j.1521-0391.2012.00256.x>
- [34] Roman, P.M. (1990). Seventh dimension, A new component is added to the EAP 'core technology.' *Employee Assistance*, 2, 8-9.
- [35] Hargrave, G.E., & Hiatt, D. (2005). The EAP treatment of depressed employees: Implications for return on investment. *Employee Assistance Quarterly*, 19(4), 39-49. doi:10.1300/J022v19n04_03
- [36] Hargrave, G.E., Hiatt, D., Alexander, R., & Shaffer, I.A. (2008). EAP treatment impact on presenteeism and absenteeism: Implications for return on investment. *Journal of Workplace Behavioral Health*, 23(3), 283-293. <https://doi.org/10.1080/15555240802242999>
- [37] Osilla, K.C., dela Cruz, E., Miles, J.N.V., Zellmer, S., Watkins, K., Larimer, M.E., & Marlatt, G.A. (2010). Exploring productivity outcomes from a brief intervention for at-risk drinking in an employee assistance program. *Addictive Behaviors*, 35, 194-200. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2821832/>
- [38] Lerner, D., Adler, D.A., Rogers, W.H., Chang, H., Greenhill, A., Cymerman, E., & Azocar, F. (2015). A randomized clinical trial of a telephone depression intervention to reduce employee presenteeism and absenteeism. *Psychiatric Services*, 66, 570-577. <https://ps.psychiatryonline.org/doi/10.1176/appi.ps.201400350>
- [39] Richmond, M.K., Shepherd, J.L., Pampel, F.C., Wood, R.C. Reimann, B., & Fischer, L. (2014). Associations between substance use, depression, and work outcomes: An evaluation study of screening and brief intervention in a large employee assistance program. *Journal of Workplace Behavioral Health*, 29, 1-18. <https://doi.org/10.1080/15555240.2014.866470>
- [40] Richmond, M.K., Pampel, F.C., Wood, R.C., & Nunes, A.P. (2015). The impact of employee assistance services on workplace outcomes: Results of a prospective, quasi-experimental study. *Journal of Occupational Health Psychology*, 22, 170-179. <https://doi.org/10.1037/ocp0000018>
- [41] Richmond, M.K., Pampel, F.C., Wood, R.C., & Nunes, A.P. (2016). Impact of employee assistance services on depression, anxiety, and risky alcohol use. *Journal of Occupational and Environmental Medicine*, 58(7), 641-650. doi:10.1097/JOM.0000000000000744
- [42] Milot, M. (2019). The impact of a Canadian external employee assistance program on mental health and workplace functioning: Findings from a prospective quasi-experimental study. *Journal of Workplace Behavioral Health*, 34(3), 167-191. <https://doi.org/10.1080/15555240.2019.1609978>
- [43] Maeng, D., Cornell, A.E., & Nasra, G.S. (2019). Utilization of an employee behavioral health program and its effects on outcomes for depression and anxiety disorders. *Journal of Occupational and Environmental Medicine*, 61(10), 812-817. doi:10.1097/JOM.0000000000001678
- [44] Schneider, R.A., Grasso, J.R., Chen, S.Y., Chen, C., Reilly, E.D., & Kocher, B. (2020). Beyond the lab, empirically supported treatments in the real world. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.01969>
- [45] Lungu, A., Jun, J.J., Azarmanesh, O., Leykin, Y., & Chen, C.E.-J. (2020). Blended care-cognitive behavioral therapy for depression and anxiety in real-world settings, pragmatic retrospective study. *Journal of Medical Internet Research*, 22, e18723. doi:10.2196/18723

- [46] Bondar, J., Morrow, C.B., Gueorguieva, R., Brown, M., Hawrilenko, M., Krystal, J.H., Corlett, P.R., & Chekroud, A.M. (2022). Clinical and financial outcomes associated with a workplace mental health program before and during the COVID-19 pandemic. *JAMA Network Open*, 5(6):e2216349 doi:10.1001/jamanetworkopen.2022.16349
- [47] Andrews, G., Basu, A., Cuijpers, P., Craske, M.G., McEvoy, P., English, C.L., et al. (2018). Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. *Journal of Anxiety Disorders*, 55, 70-78. doi:10.1016/j.janxdis.2018.01.001
- [48] Lattie, E.G., Adkins, E.C., Winquist, N., Stiles-Shields, C., Wafford, Q.E., & Graham, A.K. (2019). Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: Systematic review. *Journal of Medical Internet Research*, 21(7):e12869. doi:10.2196/12869
- [49] Attridge, M.A., Morfit, R.C., Roseborough, D.J., & Jones, E.R. (2020). Internet-based cognitive-behavioral therapy for college students with anxiety, depression, social anxiety or insomnia: Four single-group longitudinal studies of archival commercial data and replication of employee user study. *JMIR – Journal Formative Research*, 4(7), e17712 doi:10.2196/17712
- [50] Jacobson, N.S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12-19. <https://doi.org/10.1037/0022-006X.59.1.12>
- [51] Hageman, W.J., & Arrindell, W.A. (1993). A further refinement of the reliable change (RC) index by improving the pre-post difference score: Introducing RCI. *Behavior Research and Therapy*, 31(7), 693-700. [https://doi.org/10.1016/0005-7967\(93\)90122-BW](https://doi.org/10.1016/0005-7967(93)90122-BW)
- [52] Gyani, A., Shafar, R., Layard, R., & Clark, D.M. (2013). Enhancing recovery rates: Lessons from year one of IAPT. *Behavioural Research and Therapy*, 51(9), 597-606. <https://doi.org/10.1016/j.brat.2013.06.004>
- [53] McMillan, D., Gilbody, S., & Richards, D. (2010). Defining successful treatment outcome in depression using the PHQ-9: a comparison of methods. *Journal of Affective Disorders*, 127(1-3), 122-129. <https://doi.org/10.1016/j.jad.2010.04.030>
- [54] Liu, F.F., & Adrian, M.C. (2019). Is treatment working? Detecting real change in the treatment of child and adolescent depression. *Journal of the American Academy of Child & Adolescent Psychiatry*, 58(12), 1157-1164. <https://doi.org/10.1016/j.jaac.2019.02.011>
- [55] Attridge, M. (2020). *Workplace Outcome Suite (WOS) annual report 2020: Part 1 - Decade of data on EAP counseling reveals prominence of presenteeism* [White paper]. Morneau Shepell. <http://hdl.handle.net/10713/13758>
- [56] Attridge, M., Sharar, D., DeLapp, G., & Veder, B. (2018). EAP works, global results from 24,363 counseling cases with pre-post data on the Workplace Outcome Suite. *International Journal of Health and Productivity*, 10(2), 7-27. <http://hdl.handle.net/10713/8962>
- [57] Attridge M. (2020). *Workplace Outcome Suite (WOS) annual report 2020: Part 2—Profiles of work outcomes on 10 context factors of EAP counseling use* [White paper]. Morneau Shepell. <http://hdl.handle.net/10713/13759>
- [58] Attridge, M. (2011). The business case bibliography: 100 review papers on the workplace value of mental health, addiction and EAP services. *EASNA Research Notes, Vol. 2, No. 4*. <http://hdl.handle.net/10713/5129>
- [59] Masi, D.A. (1994). *Evaluating your employee assistance and managed behavioral care program*. Performance Resources Press.
- [60] Feit, M.D., & Holosko, M.J. (Eds.). (2012). *Evaluation of employee assistance programs*. Routledge. ISBN 0-86656-741-0
- [61] Blum, T., & Roman, P. (1995). *Cost-effectiveness and preventive implications of employee assistance programs* [White paper]. U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services. <https://www.ojp.gov/pdffiles1/Digitization/160889NCJRS.pdf>
- [62] McLeod, J. (2008). *Counselling in the workplace: A comprehensive review of the research evidence. 2nd edition* [White paper]. Lutterworth: BACP. <https://www.bacp.co.uk/media/1974/bacp-counselling-in-workplace-systematic-review.pdf>
- [63] McLeod, J. (2010). The effectiveness of workplace counselling: A systematic review. *Counsel Psychotherapy Research*, 10(4), 238–248. <https://doi.org/10.1080/14733145.2010.485688>
- [64] Martin, A., Sanderson, K., Cocker, F. (2009). Meta-analysis of the effects of health promotion intervention in the workplace on depression and anxiety symptoms. *Scandinavian Journal of Work and Environmental Health*, 35(1), 7–18. doi:10.5271/sjweh.1295
- [65] Joseph, B., Walker, A., & Fuller-Tyszkiewicz, M. (2017). Evaluating the effectiveness of employee assistance programmes: A systematic review. *European Journal of Work and Organizational Psychology*, 27(1), 1-15. <https://doi.org/10.1080/1359432X.2017.1374245>
- [66] Csiernik, R. (2005). A review of EAP evaluation in the 1990s. *Employee Assistance Quarterly*, 19(4), 21-37. <http://hdl.handle.net/10713/4534>
- [67] Csiernik, R. (2011). The glass is filling: An examination of employee assistance program evaluations in the first decade of the new millennium. *Journal of Workplace Behavioral Health*, 26(4), 334-355. <https://doi.org/10.1080/15555240.2011.618438>
- [68] Csiernik, R., Cavell, M., & Csiernik, B. (2021). EAP evaluation 2010–2019: What do we now know? *Journal of Workplace Behavioral Health*, 36(2), 105-124. <https://doi.org/10.1080/15555240.2021.1902336>
- [69] Attridge, M. (2010). EAP cost-benefit research: 20 years after McDonnell Douglas. Part 1 of 3. *Journal of Employee Assistance*, 40(2), 14-16. <http://hdl.handle.net/10713/5153>
- [70] Attridge, M. (2010). 20 years of EAP cost-benefit research: Taking the pareto path to ROI. Part 2 of 3. *Journal of Employee Assistance*, 40(3), 12-15. <http://hdl.handle.net/10713/14935>
- [71] Attridge, M. (2010). 20 years of EAP cost-benefit research: Taking the productivity path to ROI. Part 3 of 3. *Journal of Employee Assistance*, 40(4), 8-11. <http://hdl.handle.net/10713/14903>
- [72] Attridge, M. (2013, October). *The business value of employee assistance: A review of the art and science of ROI*. Keynote address at the annual conference of the Employee Assistance Professionals Association, Phoenix, AZ. <http://hdl.handle.net/10713/8380>
- [73] Attridge, M. (2016, October). *EAP industry outcomes for employee absenteeism and presenteeism: A global research analysis* [Paper presentation]. Employee Assistance Professionals Association, Chicago, IL, United States. <http://hdl.handle.net/10713/7203>
- [74] Luthans, F., & Waldersee, R. (1989). What do we really know about EAPs?. *Human Resource Management*, 28(3), 385-401. <https://doi.org/10.1002/hrm.3930280307>
- [75] Attridge, M. (2001). Can EAPs experiment in the real world? *EAPA Exchange*, 31(2), 26-27. <http://hdl.handle.net/10713/15064>
- [76] Pompe, J.C., & Sharar, D.A. (2008). Preparing for the challenges of research. *Journal of Employee Assistance*, 38(2), 7-9. <http://hdl.handle.net/10713/5149>
- [77] Milot, M. (2020). A researcher's inside look at the impact of employee assistance programs. *IFEBP Plans & Trusts*, 38(1), 14-19. <http://hdl.handle.net/10713/14266>
- [78] CuraLinc Healthcare. (2018). *EAP Outcomes and Impact, Case Study 2018* [White paper]. Chicago, IL, United States.
- [79] CuraLinc Healthcare. (2019). *EAP Outcomes and Impact, Case Study 2019* [White paper]. Chicago, IL, United States.
- [80] CuraLinc Healthcare. (2020). *EAP Outcomes and Impact, Case Study 2020* [White paper]. Chicago, IL, United States.
- [81] CuraLinc Healthcare. (2021). *EAP Outcomes and Impact, Case Study 2021* [White paper]. Chicago, IL, United States. <https://www.eapoutcomes.com>

- [82] Kroenke, K., Spitzer, R.L., & Williams, J.B. (2001). The PHQ-9, Validity of a brief depression severity measure. *Journal of General and Internal Medicine*, 16, 606-613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- [83] Kroenke, K., & Spitzer, R.L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, 32(9), 509-515. <https://doi.org/10.3928/0048-5713-20020901-06>
- [84] Löwe, B., Unutzer, J., Callahan, C.M., Perkins, A.J., & Kroenke, K. (2004). Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. *Medical Care*, 42(12), 1194–1201. doi:10.1097/00005650-200412000-0000
- [85] Martin, A., Rief, W., Klaiberg, A., Braehler, E. (2006). Validity of the brief patient health questionnaire mood scale (PHQ-9) in the general population. *General Hospital Psychiatry*, 28(1), 71-77. <https://doi.org/10.1016/j.genhosppsych.2005.07.003>
- [86] Wang, W., Bian, Q., Zhao, Y., Li, X., Wang, W., Du, J., et al. (2014). Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry*, 36(5), 539-544. <https://doi.org/10.1016/j.genhosppsych.2014.05.021>
- [87] Kroenke, K. (2012). Enhancing the clinical utility of depression screening. *Canadian Medical Association Journal*, 184(3), 281-282. www.cmaj.ca/lookup/doi/10.1503/cmaj.110829
- [88] Manea, L., Gilbody, S., & McMillan, D. (2012). Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis. *Canadian Medical Association Journal*, 184(3), E191-E196. <https://doi.org/10.1503/cmaj.110829>
- [89] Babor, T.F., Higgins-Biddle, J.C., Saunders, J.B., & Monteiro, M.G. (2001). *The Alcohol Use Disorders Identification Test, Guidelines for use in primary health care. 2nd Edn.* (WHO Publication WHO/MSD/MSB/01.6a). World Health Organization. <https://doi.org/10.1186/1747-597X-7-42>
- [90] Sinclair, M., McRee, B., & Babor, T.F. (1992). *Evaluation of the reliability of AUDIT*. University of Connecticut School of Medicine, Alcohol Research Center (Unpublished report).
- [91] Higgins-Biddle, J.C., & Babor, T.F. (2018). A review of the Alcohol Use Disorders Identification Test (AUDIT), AUDIT-C, and USAUDIT for screening in the United States: Past issues and future directions. *American Journal of Drug and Alcohol Abuse*, 44(6), 578-586. <https://doi.org/10.1080/00952990.2018.1456545>
- [92] Erford, B.T., Sriken, J., Sherman, M.F., Hibbs, J.S., Smith, H.L., Kipper-Smith, A., & Niarhos, F. (2021). Psychometric analysis, internal structure, and measurement invariance of the Alcohol Use Disorders Identification Test (AUDIT) scores from a large university sample. *Measurement and Evaluation in Counseling and Development*, 54(3), 188-205. <https://doi.org/10.1080/07481756.2020.1770101>
- [93] Lennox, R.D., Sharar, D., Schmitz, E., & Goehner, D.B. (2010). Development and validation of the Chestnut Global Partners Workplace Outcome Suite. *Journal of Workplace Behavioral Health*, 25(2), 107-131. <https://doi.org/10.1080/15555241003760995>
- [94] Koopman, C., Pelletier, K.R., Murray, J.F., Sharda, C.E., Berger, M.L., Turpin, R.S., et al. (2002). Stanford Presenteeism Scale: Health status and employee productivity. *Journal of Occupational and Environmental Medicine*, 44(1), 14-20. <https://doi.org/10.1097/00043764-200201000-00004>
- [95] Turpin, R.S., Ozminkowski, R.J., Sharda, C.E., Collins, J.J., Berger, M.L., Billotti, G.M., et al. (2004). Reliability and validity of the Stanford Presenteeism Scale. *Journal of Occupational and Environmental Medicine*, 46(11), 1123–1133. doi:10.1097/01.jom.0000144999.35675.a0
- [96] Consortium for Mental Healthcare. (2019). *Stanford Presenteeism Scale (SPS-6)* [White paper]. Ipsos-Reid/Guarding Minds Canada. <https://www.ppcn.org/Assessments/Presenteeism/Presenteeism-Scoring-Review.pdf>
- [97] Aronsson, G., Gustafsson, K., & Dallner, M. (2000). Sick but yet at work: An empirical study of sickness presenteeism. *Journal of Epidemiology and Community Health*, 54, 502-509. <http://dx.doi.org/10.1136/jech.54.7.502>
- [98] Lohaus, D., & Habermann, W. (2019). Presenteeism: A review and research directions. *Human Resource Management Review*, 29, 43-58. <https://doi.org/10.1016/j.hrmr.2018.02.010>
- [99] Cooper, C., & Dewe, P. (2008). Well-being, absenteeism, presenteeism costs and challenges. *Occupational Medicine*, 58, 522-524. <https://doi.org/10.1093/occmed/kqn124>
- [100] Hemp, P. (2008). Presenteeism: At work - but out of it. *Harvard Business Review*, 82, 49–58. https://www.npg-rsp.ch/fileadmin/npg-rsp/Themen/Fachthemen/Hemp_2004_Presenteeism.pdf
- [101] Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, 2nd. Ed.* Lawrence Erlbaum Associates. ISBN 0-8058-0283-5
- [102] Richardson, J.T.E. (2011). Eta squared and partial eta squared as measurements of effect size in educational research. *Educational Research Review*, 6(2), 135-147. <https://doi.org/10.1016/j.edurev.2010.12.001>
- [103] American Psychological Association. (2017). APA ethical guidelines for research. http://www.sandplay.org/pdf/APA_Ethical_Guidelines_for_Research.pdf
- [104] Carpenter, J., Crutchley, P., Zilca, R.D., Schwartz, H.A., Smith, L.K., Cobb, A.M., et al. (2016). Seeing the “big” picture, big data methods for exploring relationships between usage, language, and outcome in internet intervention data. *Journal of Medical Internet Research*, 18, e241. <https://www.jmir.org/2016/8/e241/>
- [105] Lungu, A., Boone, M.S., Chen, S.Y., Chen, C.E-J., & Walsler, R.D. (2021). Effectiveness of a cognitive behavioral coaching program delivered via video in real world settings. *Telemedicine and e-Health*, 127, 47-54. <http://doi.org/10.1089/tmj.2019.0313>
- [106] Carlbring, P., Andersson, G., Cuijpers, P., Riper, H., Hedman-Lagerlöf, E. (2018). Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. *Cognitive Behaviour Therapy*, 247(1), 1-8. <https://doi.org/10.1080/16506073.2017.1401115>
- [107] Attridge M, Dickens SP. (2022). Health and work outcomes of brief counseling from an EAP in Vermont: Follow-up survey results, client satisfaction, and estimated cost savings. *SAGE Open*;12(1), 1-19. <https://doi.org/10.1177/21582440221087278>
- [108] Attridge, M. (2021, November 5). *Employee assistance programs: Trends and technology* [Virtual presentation]. Work, Stress and Health: Virtual Conference 2021, American Psychological Association & National Institutes of Occupational Health and Safety (NIOSH). <http://hdl.handle.net/10713/17890>
- [109] Couser, G.P., Nation, J.L., & Hyde, M.A. (2020). Employee assistance program response and evolution in light of COVID-19 pandemic. *Journal of Workplace Behavioral Health*, 36(3), 197-212. <https://doi.org/10.1080/15555240.2020.1821206>
- [110] Hughes, D., & Fairley, A. (2020). The COVID chronicles: An employee assistance program’s observations and responses to the pandemic. *Journal of Workplace Behavioral Health*, 36(3), 177-196. <https://doi.org/10.1080/15555240.2020.1844569>
- [111] Attridge, M., Cahill, T., Granberry, S.W., & Herlihy, P.A. (2013). The National Behavioral Consortium Industry profile of external EAP vendors. *Journal of Workplace Behavioral Health*, 28(4), 51–324. <http://dx.doi.org/10.1080/15555240.2013.845050>
- [112] Kessler, R.C., Barber, C., Beck, A., Berglund, P., Cleary, P.D., McKeenas, D, et al. (2003). The World Health Organization Health and Work Performance Questionnaire (HPQ). *Journal of Occupational and Environmental Medicine*, 45(2), 156-174. doi:10.1097/01.jom.0000052967.43131.51
- [113] Jenkins, K.R. (2014). How valid are self-reports of illness-related absence? Evidence from a university employee health management program. *Population Health Management*, 17(4), 211-217. <https://doi.org/10.1089/pop.2013.0087>
- [114] Johns, G., & Miraglia, M. (2015). The reliability, validity, and accuracy of self-reported absenteeism from work: a meta-analysis. *J Occupational Health Psychol*. 2015;20(1):1-14. <https://doi.org/10.1037/a0037754>

- [115] Despiégl, N., Danchenko, N., François, C., Lensberg, B., & Drummond, M.F. (2012). The use and performance of productivity scales to evaluate presenteeism in mood disorders. *Value in Health*, 15(8), 1148-1161. <https://doi.org/10.1016/j.jval.2012.08.2206>
- [116] Ospina, M.B., Dennett, L., Waye, A., Jacobs, P., & Thompson, A.H. (2015). A systematic review of measurement properties of instruments assessing presenteeism. *American Journal of Managed Care*, 21(2):e171-185. <https://www.ajmc.com/view/a-systematic-review-of-measurement-properties-of-instruments-assessing-presenteeism>

AUTHORS

First Author – Mark Attridge, Ph.D., M.A., President, Attridge Consulting, Inc., 1129 Cedar Lake Road South, Minneapolis, MN 55405, USA; mark@attridgeconsulting.com; ORCID = orcid.org/0000-0003-1852-2168

Second Author – David Pawlowski, M.S., Senior Vice President of Operations, CuraLinc Healthcare, 314 W. Superior Street, Suite 601, Chicago, IL 60654, USA; dpawlowski@curalinc.com

Third Author – Sean Fogarty, M.B.A., President, CuraLinc Healthcare, 314 W. Superior Street, Suite 601, Chicago, IL 60654, USA; sfogarty@curalinc.com

Correspondence Author – Dr. Mark Attridge, mark@attridgeconsulting.com; phone: 1+612-889-2398