

Research on the Tufts Be Well at Work Program for Employees With Depression: 2005–2020

Debra Lerner, M.Sc., Ph.D., David Adler, M.D., Ariella Shayani, M.D., M.P.H., William H. Rogers, Ph.D.

Objective: Although depression is a prevalent and costly health problem exacting a large toll on work productivity, interventions targeting occupational functioning are rare. This article describes the development of the Tufts Be Well at Work intervention, a brief telephonic program designed to improve occupational functioning among employees with depression and reduce depression symptom severity. Results from 15 years of research are summarized evaluating the occupational, clinical, and economic impact of Be Well at Work.

Methods: The design, methods, and results of all six Tufts Be Well at Work studies are reported. Studies included an initial workplace pilot study, two workplace randomized clinical trials (RCTs), one RCT in a health care system, and two pilot implementation studies conducted in a workplace

and in an academic medical center. RCTs compared Tufts Be Well at Work to usual care.

Results: Tufts Be Well at Work consistently and significantly improved occupational functioning, work productivity, and depression symptom severity. Employees randomly assigned to usual care experienced smaller gains. The program also delivered a positive return on investment.

Conclusions: Evidence suggests that Tufts Be Well at Work is an effective intervention for improving occupational and clinical functioning. Its relatively low cost and its impact on work productivity contribute to its positive economic impact.

Psychiatric Services 2021; 00:1–10; doi: 10.1176/appi.ps.202000762

In 2011, our *Psychiatric Services* editorial urged researchers, clinicians, and employers to help depressed employees participate fully and effectively in the labor market (1). This recommendation was based on research that found high rates of depression-related job loss, absenteeism, and presenteeism (at-work functional difficulties) and associated productivity costs (2–8). In the United States, depression is one of the most prevalent and costly health problems (9). One in eight working-age adults is estimated to be clinically depressed, and half will experience a recurrence within 1 year of remission (10). Access to high-quality care has been constrained by gaps in insurance coverage, provider supply, physician practice patterns, stigma, and lack of consumer knowledge of depression and its treatment (11, 12).

The U.S. direct cost of care for commercially insured persons with depression is at least \$98.9 billion annually (13), even though only 50% of adults with depression are diagnosed and under half of those diagnosed receive guideline-concordant care (1, 14). At the societal level, the indirect costs of depression's morbidity and mortality, costs related to lost productivity, have been estimated to be \$111.7 billion annually in the United States (in 2010 dollars) (13). This extraordinary cost is separate from the economic burden at

the employer and employee levels. For example, it has been estimated that a company with 10,000 employees and \$70 million in profits will accrue depression costs of \$17.2

HIGHLIGHTS

- Tufts Be Well at Work is a new short-term, low-cost, work-focused, telephone-based treatment program designed to improve the occupational functioning of working people with depression.
- This work-focused intervention has four integrated components, each addressing a specific barrier to recovering occupational functioning: private electronic screening, work coaching and modification strategies, work-focused cognitive-behavioral therapy, and psychoeducational and care coordination interventions.
- Tufts Be Well at Work significantly improved occupational functioning, work productivity, and depression symptom severity and showed a positive return on investment.
- Tufts Be Well at Work may be implemented in the workplace or in health care systems and aligns well with the increased use of telehealth services.

million (25% of profit), including just under \$10 million in absenteeism and presenteeism (12). Individual employees, blue collar and white collar, are also paying a price, although not well documented, through lost wages and benefits, diminished opportunities for career advancement, and economic insecurity and their consequences in terms of the quality of life of employees and their families.

The aim of this article is to describe the research and development process for the Tufts Be Well at Work program, an intervention for improving the occupational functioning of employees with depression and reducing productivity loss. In this article, we organize the evidence regarding the program's effectiveness and cost and discuss progress and remaining challenges. The program was intended to fill a gap in care that mainly consisted of (and continues to consist of) medical care options to reduce depression symptoms (e.g., antidepressants and psychotherapy) and employer-sponsored programs, such as employee assistance programs (EAPs). Over 15 years, with the help of federal grants, the program was developed and tested in an initial pilot study (15), three randomized clinical trials (RCTs) (7, 16, 17), and two pilot implementation studies (no publications).

Clinical research on depression treatment has been focused on addressing problems of access and quality, with the goal of improving depression symptoms. Relatively few clinical research studies have considered work outcomes as the primary endpoint (18) (exceptions include a few high-quality care approaches, such as the collaborative care model, but these have been slow to disseminate [19–21]). Similarly, research on the efficacy of employer-sponsored options, such as EAPs for employees with depression, is sparse, and average annual EAP use rates remain relatively low (2%–3% of employee populations) (22). Other employer- or payer-sponsored services, such as depression screening, resiliency training, case management for complex health conditions, and disability or absence management programs, vary considerably, and there is limited evidence to support effects on depression or its work outcomes (23).

Since approximately 2014, almost a decade after we began our intervention research, several systematic reviews and meta-analyses have been published to determine the efficacy of interventions, some delivered through the workplace, on the clinical or work outcomes of adults with depression. A systematic review by Lee and colleagues (24) of 17 RCTs testing the efficacy of different antidepressants on work functioning and work absences found positive effects. Another meta-systematic review conducted by Joyce and colleagues (25) included workplace interventions of any duration for employees with depression or anxiety, most of which involved the use of cognitive-behavioral therapy (CBT), using a variety of approaches. CBT reduced symptoms but did not have a significant impact on work outcomes unless it was combined with other approaches, such as problem-solving therapy.

Tan and colleagues (26) implemented a meta-analysis of nine RCTs testing universal workplace programs for

depression prevention, most of which involved CBT, and found a small positive effect on depression symptoms. A Cochrane Collaborative systematic review, updated in 2020 (27), investigated the effects of clinical and work-directed interventions aimed at reducing disability among employees with depression. In 45 studies with more than 12,000 individuals, they found moderate evidence for the efficacy of interventions that combined clinical and work-directed strategies for reducing the number of sickness absence days and functioning. They noted gaps in evidence for longer-term effects and the need for studies to identify which specific combinations of interventions performed best. Our three RCTs were included in the clinical or work-directed studies group (7, 16, 17).

THE TUFTS BE WELL AT WORK PROGRAM

Tufts Be Well at Work is aimed at addressing this large unmet need for services with integrated and tailored mental health and vocational strategies. Although the program gives primacy to occupational functioning, the knowledge and skills that employees gain often generalize to other life domains.

Tufts Be Well at Work is a telephone-based program and, therefore, can be provided by employers; service suppliers, such as contracted EAPs; behavioral health provider organizations; health insurers; and health care systems. Care is provided by counselors who have a master's degree or higher during eight approximately 50-minute biweekly telephone sessions. The total duration is 16 weeks. Although the program is compatible with videoconferencing, we have tested the telephonic version only.

The program is supported by an electronic data system that is public facing (for employee screening, enrollment, informed consent, and data collection) and counselor facing (for scheduling, care documentation, standardized assessment, self-management planning, monitoring of counselor fidelity to the protocol, clinical supervision, and outcome reporting). The websites are privacy protected, password protected, and HIPAA compliant.

The program has four integrated components, each addressing a specific barrier to recovering occupational functioning. They are private electronic screening, accessible from a computer or mobile device to improve participation; work coaching and modification strategies targeting work-related barriers to functioning; work-focused CBT strategies to address maladaptive thoughts, feelings, and behaviors interfering with working; and psychoeducational and care coordination interventions to strengthen linkages to primary care physicians or other providers. The program also makes use of motivational interviewing, monthly reassessment, and self-management of depression principles.

Screening uses the self-administered nine-item Patient Health Questionnaire (PHQ-9) (28, 29) to assess depression symptoms and the Work Limitations Questionnaire (WLQ)

(30, 31) to assess deficits in occupational functioning. Both employee and counselor receive the results. Work coaching and modification addresses barriers to effective functioning through acquiring new work strategies. The intervention includes a detailed discussion of the employee's job demands, job control, stressors, and supports. With methods applied in vocational rehabilitation, supported employment, and management, the counselor and employee develop a work modification plan. This plan, which is designed to be self-administered and does not require employer involvement, promotes improved functioning through adopting changes in work behavior (e.g., new time management techniques), work routines (e.g., collaborating versus staying isolated), or the work environment (e.g., reducing distractions).

Work-focused CBT addresses psychological barriers to functional improvement, promoting strategies to change behaviors and cognitions that interfere with effective functioning (20, 32). Employees learn to identify the maladaptive patterns of thinking, feelings, and acting that are eroding work functioning and to substitute more adaptive habits. A series of work-related experiments, customized to the employee, are carried out.

The program also employs some of the principles of high-quality care associated with the more extensive collaborative care model (18, 33). This core component supports patient activation through motivational interviewing, psychoeducation, care coordination, self-management support, and monthly assessments. With the employee's permission, monthly assessments of symptoms and occupational functioning are sent to the employee's regular care provider. More detailed description of the intervention is available elsewhere (3, 7, 15–17).

TUFTS BE WELL AT WORK STUDIES DESIGNS AND METHODS

Over the course of 15 years, we have conducted six studies (Tables 1–3). (A table in an online supplement to this article presents further details.) These six studies were conducted in different settings and study populations, partly by design and partly opportunistically. Because we are interested in presenting both the journey and the results, we report on each study, including the nuances of design and methods, results, and conclusions, rather than conducting a meta-analysis of these studies. A recent and comprehensive systematic review, which includes our three RCTs, is available (27).

The primary objective of each of the six studies was to test the effectiveness of the program and secondarily to determine whether a return on investment (ROI) was obtained. Our research objectives, as well as the study designs and methods, reflect our original vision of the intervention as a program for employees delivered through the workplace directly by employers or their service providers. Because employers as well as employees were considered key stakeholders in the outcome of the research, most of the

study sites were connected to the workplace (e.g., private-sector companies and public-sector government employers and benefits administrators). Over time, as health care systems diversified their services and telephonic care became more common, we conducted two studies (studies 4 [17] and 6) within health care systems. By 2020, we had conducted one relatively large multisite, multiemployer RCT (study 3 [16]); three single-employer studies, each involving multiple decentralized work locations (studies 1 [15], 2 [7], and 5); and two single-site studies (studies 4 [17] and 6) (Table 1 presents a detailed overview).

From the outset, we were committed to both preserving the applied, real-world nature of our research, which was important to employer stakeholders, and maintaining scientific rigor, and studies were designed to have a high degree of internal validity. We placed a priority on high-quality, uniform methods of screening, randomly assigning, and sampling employees, as well as data collection, measurement, and statistical analysis.

Each study included employed adults (based in the United States) who met criteria for major depression or dysthymia or (with the adoption of *DSM-5*) persistent depressive disorder, had productivity loss of 5% or higher in the past 2 weeks, could read and understand English, worked at least 15 hours per week, and had no history of psychosis or bipolar disorder. Five of the six studies had a minimum age requirement of 18 (study 3 enrolled middle-aged and older employees). Eligibility criteria related to health plan or source of care were included in three studies (studies 4–6).

In five of the six studies, screening and eligibility procedures were fully web based and self-administered; study 4 required in-person and telephone screening. Several employers had specific preferences related to the intensity, timing, frequency, and content of study advertising and the use of incentives—some required us to offer screening to all employees rather than directing advertising to those with symptoms. Such open screening increased the volume of website hits and decreased site eligibility rates (Table 2) while increasing variation in rates across studies. Ultimately, we completed 26,356 screenings, 2,251 individuals met study eligibility criteria, and 1,037 enrolled. The enrollment rate for the RCTs ranged from 35% to 88%. The lower rate reflects the variety in the study involving 24 workplaces (study 3). The higher rate reflects the benefits of enrolling in an academic medical center (study 4).

Among the studies, similar screening and data collection tools were used. Every study had a core set of screening, baseline, and follow-up questionnaires, which included validated and internationally used tools, such as the PHQ-9 for depression, the WLQ for presenteeism, and the WLQ time loss module for absenteeism. The timing of study observations was consistent with baseline at preintervention and follow-up at immediately postintervention (month 4). In the study at the Veterans Affairs medical center (VAMC) (study 4 [17]), a second follow-up at study month 8 was included,

TABLE 1. Overview of studies of Tufts Be Well at Work program

Study characteristic	Study 1: Global security company (U.S. only) (15)	Study 2: State government employer (7)	Study 3: Health care services group and individual employers (16)	Study 4: VA medical center (VAMC) behavioral health lab (17)	Study 5: State government employer	Study 6: Academic medical center
Site	Private employer (3 locations) April 2006–March 2007 (12)	Government employer October 2007–April 2008 (7)	Multiple employers (24 employers) ^a 2010–2013 (36)	Outpatient health clinic 2014–2018 (39)	Government employer (2 divisions) August 2015–October 2016 (14)	Private employer April 2019–November 2019 (8)
Design	Observational study with external comparator	2-arm randomized controlled trial (RCT) with 2:1 ratio	2-arm RCT with 1:1 ratio	2-arm RCT with 1:1 ratio	Pre-post	Pre-post
Screening method	Web-based, privacy-protected study website 8 biweekly 50-minute phone sessions	Web-based, privacy-protected study website 8 biweekly 50-minute phone sessions	Web-based, privacy-protected study website 8 biweekly 50-minute phone sessions	Phone prescreening and in-person screening interview 8 biweekly 50-minute phone sessions plus 1 phone session visit 4 months later	Web-based, privacy-protected, study website 8 biweekly 50-min phone sessions	Web-based, privacy-protected, study website 8 biweekly 50-min phone sessions
Experimental intervention						
Intervention provider qualifications	Masters-level, licensed employee assistance program (EAP) counselors	Masters-level, licensed EAP counselors	Masters-level, licensed EAP counselors	VAMC doctoral-level psychologists	Master's-level licensed counselor	Masters-level licensed counselor
Comparator ^b	Secondary data on usual care	Usual care	Usual care	VAMC integrated behavioral care	Usual care	Usual care
Eligibility criteria	Age ≥18	18–62	≥45	≥18	≥18	≥18
Employment	≥15 hours per week	≥15 hours per week	≥15 hours per week	≥15 hours per week, in job ≥6 months	≥15 hours per week	≥15 hours per week
Other	None	None	None	VA health plan enrollment	Employer-sponsored plan enrollment >1 year	Employer-sponsored plan enrollment
Exclusion	Retiring in ≤2 years, current or pending disability claim, pregnant, severe physical limitations, non-English speaking or reading	Retiring in ≤2 years, current disability claim, current alcohol abuse or dependence or drug abuse; pregnant or 6 months postpartum, schizophrenia, bipolar disorder, diagnosed as having ≥12 potentially disabling conditions, recent bereavement, non-English speaking or reading	Mania or bipolar disorder, psychosis, current alcohol abuse or dependence, severe physical limitations, non-English speaking or reading	Bipolar disorder, psychosis, planned maternity leave, non-English speaking or reading	Current alcohol abuse or dependence, mania, bipolar disorder, psychosis	Current alcohol abuse or dependence, mania, bipolar disorder, psychosis

continued

TABLE 1, continued

Study characteristic	Study 1: Global security company (U.S. only) (15)	Study 2: State government employer (7)	Study 3: Health care services group and individual employers (16)	Study 4: VA medical center (VAMC) behavioral health lab (17)	Study 5: State government employer	Study 6: Academic medical center
WLQ productivity loss score in past 2 weeks ^c	≥5%	≥5%	≥5%	≥5%	≥5%	≥4%
PHQ-9 symptom severity score ^d	≥10	≥5	≥5	≥5	≥5	≥7
Outcome metric, primary (P) or secondary (S)						
WLQ at-work productivity loss score ^e	P	P	P	P	P	P
Limitations in ability to perform job tasks ^f	P	P	P	P	P	P
Productivity loss due to absence ^g	P	P	P	P	P	P
PHQ-9 symptom severity score ^d	P	S	S	P	P	P
Return on investment ^h	S	S	S	S	S	S
Postbaseline timing of follow-up measurement						
Monthly					✓	
4 months		✓	✓	✓	✓	
6 months	✓					
8 months				✓		✓
12 months	✓					
18 months	✓					
Data collection modality	Online self-report	Online self-report	Online self-report	Interviewer- administered phone survey	Online self-report plus phone administration of surveys for nonresponders	Online self-report

^a Some study sites included grouped customers of the health services organization.

^b Usual care refers to any care obtained by the study participant, which may have included employer-sponsored health care, EAP or other services, other noncovered services, or no care. Usual care was not paid for by the study.

^c WLQ, Work Limitations Questionnaire. Possible scores range from 0% to 25%, with higher scores indicating greater productivity loss.

^d Possible scores on the nine-item Patient Health Questionnaire (PHQ-9) range from 0 to 27, with higher scores indicating greater severity.

^e At-work productivity loss reflects the estimated percentage difference in at-work productivity between a person (or group) completing the WLQ and an external benchmark of healthy workers.

^f As measured by the WLQ. Measure of the percentage of time the person was limited in the past 2 weeks in ability to perform job tasks (for example, time management, physical tasks, mental and interpersonal tasks, output tasks).

^g Based on responses to the WLQ time loss module.

^h Dollars saved: cost minus benefit divided by cost.

TABLE 2. Baseline characteristics of Tufts Be Well at Work study populations

Characteristic	Study 1: Global security company (U.S. only) (15)		Study 2: State government employer (7)		Study 3: Health care services group and individual employers (16)		Study 4: VA medical center (VAMC) behavioral health lab (17)		Study 5: State government employer		Study 6: Academic medical center	
	N	%	N	%	N	%	N	%	N	%	N	%
Enrollment data												
N screened ^a	1,392		1,525		18,102		367		4,670		300	
Eligible	225	16	193	12.7	1,227	6.8	287	78	294	6	25	8.3
Eligible and enrolled in study	93	41	79	41	431	35	253	88	156	53	25	100
Enrolled in intervention	86	92	52	66	217	50	139	55	156	100	25	100
Enrolled in usual care	na		27	34	214	50	114	45	na		na	
Included in analysis	76	82	72	91	380	88	211	83	87	56	19	76
Demographic and general medical												
Age (M±SD)	44.6±9.1		45.6±9.4		54.7±6.1		45.7±11.6		48.0±10		36.1	
Female	47	50	62	79	310	72	35	14	136	87	22	88
Non-Hispanic White	86	93	78	99	379	88	134	53	126	81	47	56
Married or cohabitating	61	66	37	47	224	52	na	na	89	57	10	40
Bachelor's degree or higher					293	68	71	28	97	62	17	68
Annual income (median dollars and IQR) ^b	64K	35K	31K	21K	63K	40K	45K	40K	na	58K	na	60K
≥1 comorbid general medical conditions ^c	77	83	62	79	375	87	245	97	151	97	na	
Depression symptoms												
PHQ-9 symptom severity score (M±SD) ^d	13.5±4.4		12.8±5.2		14.3±5.0		14.5±4.8		12.3±5.2		17.5±5.0	
Ever taken an antidepressant	na		45	57	336	78	68	27	112	72	na	
Taking antidepressant	27	29	43	54	241	56	22	8.7	45	29	6	24
Presenteeism												
% of at-work productivity loss in past 2 weeks (M±SD) ^e	9.7±4.3		10.2±4.1		10.3±4.4		12.3±4.7		9.3±4.5		10.8±4.1	
% of time with at-work limitations (M±SD) ^f												
Time management	45.3±23.3		44.9±19.7		42.7±22.0		52.0±23.5		41.0±22.6		48.7±20.2	
Physical job tasks	15.4±17.4		21.5±21.5		22.3±20.2		41.3±24.7		26.0±24.1		21.1±17.0	
Mental or interpersonal job tasks	35.3±16.3		37.7±15.3		38.1±17.3		46.8±21.4		32.4±17.8		39.0±15.4	
Output tasks	39.1±22.2		40.1±22.9		42.3±23.5		46.1±24.3		36.2±23.2		39.5±24.2	
Absenteeism ^g												
Absence days in past 2 weeks (M±SD)	2.7±1.2		1.5±1.6		1.6±2.2		1.8±2.4		2.0±2.1		1.12±1.9	
% productivity loss due to absence (M±SD)	25.9±12.8		15.0±14.6		14.6±18.8		18.6±22.4		21.3±22.5		na	

^a May include duplicates because individuals were permitted to take the screener more than once.^b IQR, interquartile range for the median value.^c Chosen from a list of chronic conditions, which varied in number enumerated depending on the study. Each list included an "other" response option, with space for inserting names of conditions not included in the list.^d Possible scores on the nine-item Patient Health Questionnaire (PHQ-9) range from 0 to 27, with higher scores indicating more severe depressive symptoms.^e Based on responses to the Work Limitations Questionnaire (WLO). At-work productivity loss reflects the estimated percentage difference in at-work productivity between a person (or group) completing the WLO and an external benchmark sample of healthy workers. Possible scores range from 0% to 25%, with higher scores indicating greater productivity loss.^f Based on WLO scale score indicating the percentage of time the person was limited in the past 2 weeks in ability to perform job tasks (for example, time management). Possible scale scores range from 0 to 100, with higher scores indicating a greater percentage of time limited.^g Based on responses to the WLO absence module. Productivity loss is the mean percentage of hours missed in the past 2 weeks divided by the total number of hours usually worked in that period. Possible days missed range from 0 to 14. Possible percentage productivity loss due to absence ranges from 0 to 100, with higher scores indicating greater productivity loss.

and all follow-up data were collected by telephone interviewers blinded to participant treatment group.

INTERVENTION PROTOCOL

For the RCTs (studies 2–4 [7, 16, 17]), we used simple random sampling generally, assigning participants to the intervention or to usual care. Prior to intervening, we tested for treatment group equivalency on multiple clinical and demographic variables at baseline (preintervention). We also went the extra step of testing for group equivalence on occupational variables, such as psychosocial job demands and control, by using validated assessment metrics (34). Bias related to randomization, enrollment, and attrition was explored in detail throughout each study. In all studies, counselors were not blinded to treatment group assignment. To minimize treatment contamination, program counselors did not provide usual care.

In all RCTs, all participants could participate in usual care, although the nature of the usual care was not specified in advance and not paid for by the study. Subsequently, across studies, the usual care (whatever usual employer-sponsored health care, EAP, or other services were offered by the employer health insurer and EAP) obtained by study enrollees was highly variable, even within study sites. This variability reflects the realities of service access and use and employee preferences.

Five of the six studies used identical intervention models and each of the four components. In the VAMC study (17), care coordination was already available, and only work coaching and modification and CBT were provided as the intervention. Consistent with our general approach to real-world testing, three RCTs (studies 2–4 [7, 16, 17]) and one pilot implementation study (study 5) involved collaborations with the EAP contracted to the main collaborating organization. For example, study 2 involved the employer's EAP in providing counselors, whereas in study 3, which involved many employers, Optum EAP supplied the study counselors.

Every study used our counselor-facing website to standardize intervention procedures. Because the counselors have some discretion in choosing specific intervention strategies, such as designing experiments for the CBT and work coaching portions, the website included several reference materials, including "Tools and Tips" containing strategies to improve functioning.

TUFTS BE WELL AT WORK STUDY RESULTS

Except for the VAMC study (study 4), most participants were female (Table 2). Differences in age, race-ethnicity, and income were also reflective of the composition of the study sites. Although not shown, many different occupations were represented across the six studies.

Mean baseline PHQ-9 scores in the six studies ranged from 12.3 to 17.5 (on a scale of 0–27), indicating moderate

levels of depression severity. In the RCTs, 25%–44% of participants had major depressive disorder only, 16%–50% had dysthymia or persistent depressive disorder, and 25%–40% had double depression (i.e., both). More than 50% had a history of antidepressant treatment, except in the VAMC sample (study 4), in which the rate was 27% (Table 2).

Regarding baseline occupational functioning (Table 2), the mean amount of time in the prior 2 weeks that enrollees had difficulty with time management ranged from 41.0% to 52.0%. Limitations performing physical job tasks ranged from 15.4% to 41.3%. Limitations in performing mental and interpersonal job tasks ranged from 32.4% to 46.8%. Difficulty performing output-related tasks ranged from 36.2% to 46.1%. Assuming 40 weekly work hours (80 hours over 2 weeks), occupational functioning was impaired approximately 32 to 40 hours in a 2-week period. These scores are typical of those we have observed in our other studies. On the basis of the WLQ productivity loss score algorithm, these scores translate into an estimated productivity loss at work of between 9.3% and 12.3% across studies. The mean number of days missed in the 2 weeks prior to baseline ranged from 1.1 to 2.7 days. Related productivity losses were between 15.0% and 25.9%.

Table 3 highlights the change from baseline to follow-up in outcomes and, for the RCTs, the degree to which the change differed between the intervention and the usual care groups. The specific statistical techniques included multiple regression, analysis of covariance to test for group differences, and mixed models for confirmation of effects and reporting of group differences. The non-RCTs analyzed changes from baseline to follow-up without adjustment (see table in online supplement).

In the RCTs, the adjusted mean changes in the percentage of at-work productivity loss in the intervention group were –3.5 for study 2, –4.5 for study 3, and –2.1 for study 4. In the usual care group, the mean changes in this primary endpoint were –0.3, –1.4, and –0.1, respectively (Table 3). In each RCT, the between-group differences in the changes were statistically significant. In the intervention group, the productivity gains were between 2% and 6%. In the usual care group, gains were minimal (0.3% to 1.4%).

For depression symptom severity, the adjusted mean changes in the PHQ-9 score for the intervention group were –5.4, –7.3, and –3.8 in each RCT, and in the usual care group, they were 0.6, –3.7, and –1.6, respectively (Table 3). Each of the differences in changes was statistically significant in the intervention group, demonstrating substantial improvement in PHQ-9 symptom severity (4–10 points).

In the non-RCTs (the original pilot and two implementation studies), the unadjusted mean changes in the percentage of at-work productivity loss ranged from –3.0 to –6.9. The unadjusted mean changes in depression symptom severity ranged from –6.1 to –10.6 (Table 3).

In the VAMC study (study 4), we conducted a second follow-up 4 additional months from the immediate postintervention

TABLE 3. Changes from baseline to follow-up in productivity loss and depression symptom severity, by study group

Study	Design ^a	WLQ % of at-work productivity loss ^b			PHQ-9 depression symptom severity score ^c		
		Baseline (M)	Pre-post change (M)		Baseline (M)	Pre-post change (M)	
			Intervention	Usual care		Intervention	Usual care
1. Global security firm (15)	Observational	9.7	−6.4	na	15.5	−6.5	na
2. State government employer (7)	RCT	10.3	−3.5	−.3	13.1	−5.4	.6
3. Health care services group and individual employers (16)	RCT	10.2	−4.5	−1.4	14.4	−7.3	−3.7
4. VA medical center behavioral health lab (17)	RCT	12.4	−2.1	−.1	14.1	−3.8	−1.6
5. State government employer	Implementation pilot	9.3	−3.0	na	12.3	−6.1	na
6. Academic medical center	Implementation pilot	10.8	−6.9	na	17.5	−10.6	na

^a RCT, randomized controlled trial.

^b As measured by the Work Limitations Questionnaire (WLQ), at-work productivity loss reflects the estimated percentage difference in at-work productivity between a person (or group) completing the WLQ and an external benchmark sample of healthy workers. Possible scores range from 0% to 25%, with higher scores indicating greater productivity loss.

^c Possible scores on the nine-item Patient Health Questionnaire (PHQ-9) range from 0 to 27, with higher scores indicating more severe depressive symptoms.

period. Using mixed models, we found that the intervention group's initial gain in at-work productivity loss at first follow-up did not erode at second follow-up (mean difference between the two follow-ups, −0.5; 95% confidence interval [CI]=−1.9 to 0.9, $p=0.46$). No significant loss was noted in the initial improvement in depression severity (mean difference between the two follow-ups, 0.6; 95% CI, −0.9 to 2.1, $p=0.44$).

As to whether the changes from baseline to 4 months are meaningful clinically or financially, several guideposts are useful for putting these into perspective. The improvement in symptom severity observed in the intervention groups in all studies (VAMC study 4 had smaller gains) is comparable to the reductions seen in antidepressant RCT trials that compared treatment to placebo (18, 35). For the presenteeism and absenteeism productivity parameters, a 1–percentage point productivity improvement is meaningful at group, company, and national levels. For example, a 2%–3% change in output nationally is regarded as an important indicator (36). On an employer level, if 6% of employees in a 10,000-person company with mean annual earnings of \$45,000 were to have depression, a 1% improvement in productivity would generate productivity savings of \$450 per employee with depression, or \$270,000 companywide.

Finally, to assess ROI, we computed the intervention's marginal benefits over those of usual care, minus the marginal program cost, divided by the marginal program cost. ROI is annualized. Marginal cost excluded website maintenance and research administrative costs. The benefit was the monetized value of the reduction in productivity loss due to presenteeism and absenteeism. Incremental benefits

in productivity loss were then multiplied by median earnings to convert savings into dollars. In the three RCTs, the ROI was 5.4:1, 5.2:1, and 1.6:1. The lowest ROI was from the VAMC study (study 4), which had a smaller effect on productivity loss because of work absences and a relatively higher counselor cost (see table in online supplement). The bottom line is that the Tufts Be Well at Work program was both effective and worth the investment, particularly from the employer perspective, with a cost savings range of \$671–\$979 per treated employee.

DISCUSSION

Much of the focus around psychiatric services delivery has been on individuals with identified mental disorders treated in the mental health system. We used a population-based approach, including universal screening for depression, to address the greater need of employed individuals, who seldom see mental health professionals, through a work-focused intervention (37). Our experience with the entire set of studies suggests that the Tufts Be Well at Work program is practical to implement and certainly has a place in the health care delivery system, which increasingly relies on telemental health services. As reviewed in the recent Cochrane Collaborative systematic review (27), results from the three RCTs suggest that the intervention is effective for employees with depression and deficits in occupational functioning and provides an ROI both in the workplace and in a health care system. We also have some evidence, although limited, that effects remained after the program ended. We are optimistic that we will have further studies, which will facilitate further evidence synthesis and economic analysis.

From the very beginning, we considered the importance of translating the Tufts Be Well at Work program to the workplace and working people. We were committed to moving care outside the clinical setting, making it accessible to busy people who may not have been willing or able to leave work and go to appointments. We were also aware that employer adoption of the intervention depended on making both a clinical and a business case. The business case is mainly based on cost considerations, although both payers and purchasers are becoming more interested in evidence of value that includes both treatment benefits and costs for employees and their employed dependents.

In the years between the first and last study, the employee health services landscape and patient care have changed. Workplace service offerings have become quite diversified, and companies frequently have contracts with multiple vendors, whose services have become more specialized. As a result, the EAP is no longer the only place for Tufts Be Well at Work to reach the workplace. Employers and payers have become more reliant on Internet-based mental health services, which are lower cost and relatively easy to implement and, at least in theory, offer access to care. These changes suggest that there is more opportunity for Tufts Be Well at Work to become part of a comprehensive workplace mental health care strategy. However, although many payers and purchasers can appreciate the importance of addressing occupational functioning and productivity loss, they are still bogged down by trying to reduce barriers to primary care and specialty care.

We also deliberately sought to address the concerns of a clinical audience interested in high-quality care by developing a manual-based, protocol-driven program that relies on developments in CBT, psychoeducation, patient activation, telemedicine, and, importantly, measurement-based care delivery and outcome assessment. We understood that for Tufts Be Well at Work to be disseminated and implemented, health care providers will need to appreciate both the need for such services, the effectiveness of the program for their patients, and the ease of using it. In the overburdened world of health care providers and systems generally, and during the COVID-19 pandemic specifically, this is still a challenge. However, when we started testing the program, there was no reimbursement mechanism to cover the care. Telehealth as a successful mode of behavioral treatments has slowly emerged, but until the pandemic, reimbursement for counseling was restricted by both federal and state regulations. This situation has changed, possibly making it more feasible financially to integrate Be Well at Work with patient care delivery systems.

The ability to move beyond the research arena into pilot implementation programs was strengthened by the conceptual model and study design, as well as by a primary emphasis on work outcomes as the clinical goal delivered in multiple real-world employment settings. The studies of Tufts Be Well at Work used extensively validated self-report

measures of depression and work outcomes and an accessible protocol-driven intervention with careful monitoring and documentation of treatment and systematic efforts to follow up enrollees. Limitations included the reliance on self-report measures of outcomes; lack of administrative, independent work-functioning data; and the generally brief follow-up periods that did not permit observation of whether effects are long lasting. In addition, no head-to-head comparisons of our program with other programs have been conducted. Finally, we were unable to perform a formal cost-effectiveness analysis.

CONCLUSIONS

Clinical and translational research and development of Tufts Be Well at Work was not entirely a linear process of phases, with one study building upon another through several implementation phases. By necessity, we responded to available funding and collaborative opportunities in order to advance the science and practice of functional improvement for people with depression. We have learned that we can improve occupational functioning with an evidenced-based, short-term, low-cost program and provide care that is highly convenient, acceptable, and accessible to working people. The rationale and need for this program were strong before the COVID-19 pandemic. Now, with mental distress and job insecurity skyrocketing and creating additional pressures on working people, the rationale and need are even stronger. Accelerating the dissemination and sustainable implementation of high-quality, efficient care is imperative to improve the employment outlook of adults with depression.

AUTHOR AND ARTICLE INFORMATION

Program on Health, Work and Productivity, Institute for Clinical Research and Health Policy Studies (Lerner, Adler, Rogers), and Departments of Psychiatry and Medicine (Lerner, Adler), Tufts Medical Center, Boston; Department of Psychiatry, University of Massachusetts Medical School, Worcester (Shayani). Send correspondence to Dr. Adler (dadler@tuftsmedicalcenter.org).

Drs. Lerner and Adler contributed equally to this article.

The six studies described here were supported by grant 5R01DP00101 from the Centers for Disease Control and Prevention; grants 5R34MH72735 and 5R01MH058243 from the National Institute of Mental Health; grant 5R01AG033125 from the National Institute on Aging; and Department of Veterans Affairs Merit Award 5101RX001132.

The authors acknowledge the contributions of the many collaborators on Tufts Be Well at Work studies, including John Allen, Ph.D., Francisca Azocar, Ph.D., Eugene W. Baker, Ph.D., Tom Blumenthal, L.I.C.S.W., Kevin Caskey, M.S.W., Hong Chang, Ph.D., Elina Cymerman, Ph.D., W. Dennis Derr, Ed.D., Rebecca Deusser, M.S., Andrea Dowdy, D.C., Gail Duncan, M.S.W., David Goehner, L.C.S.W., Annabel Greenhill, M.A., Amy Helstrom, Ph.D., Richard C. Hermann, M.D., Doris Hernandez, Margaret Hollins, M.S., Erin Ingram, B.A., Evette J. Ludman, Ph.D., Mercedes Lyson, Ph.D., William McPeck, M.S.W., David W. Oslin, M.D., and Pamela Thomas, M.D., M.P.H.,

Drs. Lerner, Adler, and Rogers are cofounders of Health and Productivity Sciences, a company with no current assets. Dr. Lerner reports

owning stock in Mylan and serving as a consultant for Janssen Scientific Affairs, LLC. Dr. Adler reports receipt of grant support from Janssen Pharmaceuticals. The other authors report no financial relationships with commercial interests.

Received October 15, 2020; revision received January 26, 2021; accepted February 4, 2021; published online May 7, 2021.

REFERENCES

- Lerner D, Adler DA: Improving symptoms and ability to work. *Psychiatr Serv* 2011; 62:3
- Adler DA, McLaughlin TJ, Rogers WH, et al: Job performance deficits due to depression. *Am J Psychiatry* 2006; 163:1569–1576
- Lerner D, Adler DA, Rogers WH, et al: Work performance of employees with depression: the impact of work stressors. *Am J Health Promot* 2010; 24:205–213
- Buist-Bouwman MA, Ormel J, de Graaf R, et al: Functioning after a major depressive episode: complete or incomplete recovery? *J Affect Disord* 2004; 82:363–371
- Hirschfeld RM, Dunner DL, Keitner G, et al: Does psychosocial functioning improve independent of depressive symptoms? A comparison of nefazodone, psychotherapy, and their combination. *Biol Psychiatry* 2002; 51:123–133
- Judd LL, Paulus MJ, Schettler PJ, et al: Does incomplete recovery from first lifetime major depressive episode herald a chronic course of illness? *Am J Psychiatry* 2000; 157:1501–1504
- Lerner D, Adler D, Hermann RC, et al: Impact of a work-focused intervention on the productivity and symptoms of employees with depression. *J Occup Environ Med* 2012; 54:128–135
- Trivedi MH, Morris DW, Wisniewski SR, et al: Increase in work productivity of depressed individuals with improvement in depressive symptom severity. *Am J Psychiatry* 2013; 170:633–641
- Young JQ, Kline-Simon AH, Mordecai DJ, et al: Prevalence of behavioral health disorders and associated chronic disease burden in a commercially insured health system: findings of a case-control study. *Gen Hosp Psychiatry* 2015; 37:101–108
- Marcus SC, Olfson M: National trends in the treatment for depression from 1998 to 2007. *Arch Gen Psychiatry* 2010; 67:1265–1273
- Putnam K, McKibbin L: Managing workplace depression: an untapped opportunity for occupational health professionals. *AAOHN J* 2004; 52:122–129
- Lerner D, Lyson M, Sandberg E, et al: The High Cost of Mental Disorders: Facts for Employers. Rutherford, CA, One Mind Initiative, 2017
- Greenberg PE, Fournier AA, Sisitsky T, et al: The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry* 2015; 76:155–162
- Practice Guidelines for the Psychiatric Evaluation of Adults. Washington, DC, American Psychiatric Association, 2016
- Lerner D, Adler D, Hermann RC, et al: Depression and work performance: the Work and Health Initiative Study; in *Work Accommodation and Retention in Mental Health*. Edited by Schultz IZ, Rogers ES, New York, Springer, 2011
- Lerner D, Adler DA, Rogers WH, et al: A randomized clinical trial of a telephone depression intervention to reduce employee presenteeism and absenteeism. *Psychiatr Serv* 2015; 66:570–577
- Lerner D, Adler DA, Rogers WH, et al: Effect of adding a work-focused intervention to integrated care for depression in the Veterans Health Administration: a randomized clinical trial. *JAMA Netw Open* 2020; 3:e200075
- Gilbody S, Bower P, Fletcher J, et al: Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. *Arch Intern Med* 2006; 166:2314–2321
- Unützer J, Katon W, Callahan CM, et al: Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA* 2002; 288:2836–2845
- Wang PS, Simon GE, Avorn J, et al: Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial. *JAMA* 2007; 298:1401–1411
- Schoenbaum M, Unützer J, McCaffrey D, et al: The effects of primary care depression treatment on patients' clinical status and employment. *Health Serv Res* 2002; 37:1145–1158
- Employee Benefits: The Evolution of Benefits. Alexandria, VA, Society for Human Resource Management, 2018
- Bond GR, Lerner D, Drake RE, et al: Work-Focused Interventions for Depression: Final Report. Washington, DC, US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, 2019
- Lee Y, Rosenblat JD, Lee J, et al: Efficacy of antidepressants on measures of workplace functioning in major depressive disorder: a systematic review. *J Affect Disord* 2018; 227:406–415
- Joyce S, Modini M, Christensen H, et al: Workplace interventions for common mental disorders: a systematic meta-review. *Psychol Med* 2016; 46:683–697
- Tan L, Wang MJ, Modini M, et al: Preventing the development of depression at work: a systematic review and meta-analysis of universal interventions in the workplace. *BMC Med* 2014; 12:74
- Nieuwenhuijsen K, Verbeek JH, Neumeyer-Gromen A, et al: Interventions to improve return to work in depressed people. *Cochrane Database Syst Rev* 2020; 10:CD006237
- Kroenke K, Spitzer RL, Williams JB: The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001; 16:606–613
- Kroenke K, Spitzer RL: The PHQ-9: a new depression diagnostic and severity measure. *Psychiatr Ann* 2002; 32:509–515
- Lerner D, Amick BC 3rd, Rogers WH, et al: The Work Limitations Questionnaire. *Med Care* 2001; 39:72–85
- Lerner D, Amick BC 3rd, Lee JC, et al: Relationship of employee-reported work limitations to work productivity. *Med Care* 2003; 41:649–659
- Beck J: *Cognitive Therapy: Basics and Beyond*. New York, Guilford, 1995
- Katon W, Unützer J, Wells K, et al: Collaborative depression care: history, evolution and ways to enhance dissemination and sustainability. *Gen Hosp Psychiatry* 2010; 32:456–464
- Karesek R, Theorell T: *Healthy Work Stress, Productivity and the Reconstruction of Working Life*. New York, Basic Books, 2010
- Trivedi MH, Rush AJ, Wisniewski SR, et al: Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry* 2006; 163:28–40
- Total Economy Database: Key Findings, 2019. New York, Conference Board, 2020
- Caplan LJ: NIDILRR: supporting applied research and development to improve the lives of people with mental illness. *Psychiatr Serv* 2020; 71:962–963