CADTH

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

Telehealth for the Assessment and Treatment of Depression, Post-Traumatic Stress Disorder, and Anxiety: Clinical Evidence

Service Line:	Rapid Response Service
Version:	1.0
Publication Date:	April 10, 2018
Report Length:	43 Pages

CADTH

Authors: Eugenia Palylyk-Colwell, Charlene Argáez

Cite As: Telehealth for the assessment and treatment of depression, post-traumatic stress disorder, and anxiety: clinical evidence. Ottawa: CADTH; 2018 Apr. (CADTH rapid response report: summary with critical appraisal).

ISSN: 1922-8147 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Abbreviations

BATD	behavioral activation therapy
BDI-II	Beck Depression Inventory II
BHS	Beck Hopelessness Scale
BSI-18	Brief Symptom Inventory-Short Version
CBT	cognitive behavioral therapy
CI	confidence interval
DSM-IV	Diagnostic and Statistical Manual (of Mental Disorders)-IV
GADSS	Generalized Anxiety Disorder Severity Scale
GDS	Geriatric Depression Scale
HAM-D	Hamilton Depression Rating Scale
IPT	Interpersonal psychotherapy
OR	odds ratio
PDSS	Panic Disorder Severity Scale
PE	prolonged exposure
PHQ-9	Patient Health Questionnaire-9 item
PSQI	Pittsburgh Sleep Quality Index
PTSD	post-traumatic stress disorder
RCT	randomized controlled trial
SCID	Structured Clinical Interview for the DSM-IV Axis I Disorders
SD	standard deviation
SIGH-A	Hamilton Anxiety Rating Scale
SMD	standardized mean difference
SR	systematic review
TAU	treatment as usual
T-CSCT	Telepsychiatry-based Culturally-Sensitive Collaborative Treatment

Context and Policy Issues

Telehealth encompasses a wide and rapidly evolving range of telecommunication technologies that are used to deliver health services for various conditions, populations, and settings.[1,2] For the purposes of this report, telehealth is defined as the provision of information and health care services to a specific patient by a provider in real-time using telephone, videochat, or videoconferencing technologies (e.g., Skype, FaceTime). During the telehealth sessions, the health care provider and patient can see and hear each other without being in the same room. This differs from other telehealth interventions such as on-line or computer-based therapies, smart phone applications, text messaging, or automated self-management techniques.[3-5]

Telehealth may be a viable option for patient assessment and delivery of psychotherapy to patients who are processing trauma as part of their mental health problems as well as dealing with depression, anxiety, and post-traumatic stress disorder (PTSD). In 2016, the Canadian Mental Association and the Canadian Psychiatric Association issued a joint statement on access to mental health care that included a recommendation for telehealth and e-health services to be developed as a means to ensure adequate access to mental health services.[6] In 2012, the Canadian Community Health Survey - Mental Health reported that more than 3.1 million Canadians 15 years and older experienced a major depressive episode over their lifetime and over 2.4 million experienced generalized anxiety disorder.[7] It was also reported that 476,129 Canadians had a current diagnosis of PTSD.[7] A 2013 survey of full-time regular members of the Canadian Forces revealed that, in the preceding 12 months, 8.0% had experienced a major depressive episode, 4.7% had generalized

anxiety disorder, and 5.3% had PTSD.[8] Even though public awareness about mental illness has increased substantially in recent years, ensuring access to timely and effective mental health services remains a challenge in Canada.[6]

Barriers to obtaining in-person treatment for mental health disorders includes limited accessibility to services (e.g., remote location), lack of transportation or funds, shortage of facilities and/or trained providers, stigma associated with seeking treatment, and physical mobility-related health issues.[9] Although telehealth may be a more readily accessible option, there are challenges to providing psychotherapy remotely such as establishing therapeutic alliance between provider and patient, effective management of intense patient emotions during sessions, assurance of patient safety, patient attrition, and technical challenges (e.g., reliable Internet service, image resolution, audio quality and delay).[10]

More information is needed to determine if telehealth is as clinically effective as traditional in-person psychotherapy for patients with depression, anxiety, or PTSD. There is uncertainty regarding the comparability of clinical outcomes between the treatment modalities and whether or not using telehealth influences the therapeutic process.[10] There is potential for more widespread use of telehealth to deliver psychotherapy in the future and so it is important to understand how telehealth directly compares with traditional in-person delivery of the same psychotherapeutic intervention (e.g., cognitive behavior therapy [CBT] by telephone versus in-person CBT).

The purpose of this report is to synthesize and critically appraise the available evidence on the clinical effectiveness of telepsychotherapy for the assessment and treatment of major depression, generalized anxiety disorder, and PTSD in adult patients. This report is an update of a 2015 CADTH Rapid Response Report on "Tele-medicine for Patients with Mental Health Disorders: Clinical and Cost-effectiveness".[11]

Research Questions

1. What is the clinical effectiveness of psychological assessment delivered via telehealth for patients with depression, post-traumatic stress disorder, or anxiety?
2. What is the clinical effectiveness of psychological treatment delivered via telehealth for patients with depression, post-traumatic stress disorder, or anxiety?

Key Findings

Limited evidence suggests that psychological assessment via telehealth is clinically effective in patients with depression. No relevant evidence was identified to support use of telehealth for psychological assessment in patients with anxiety or post-traumatic stress disorder.

There was concurrence among included studies that the evidence supports that psychological treatment delivered by telehealth is clinically effective and that the magnitude of the treatment effect is comparable when the same intervention is provided by telehealth or by traditional means such as in-person/same room therapy. In general, the overall risk of bias in the included studies is considered to be medium. Key limitations are the lack of patient blinding in all the included randomized controlled trials, the broad patient populations of interest, and the large number and diversity of the scales and instruments used to measure treatment effects which complicated making comparisons between studies.

Methods

Literature Search Methods

A limited literature search was conducted on key resources including Medline, PsycInfo. and PubMed, The Cochrane Library, University of York Centre for Reviews and Dissemination databases, and a focused Internet search. Methodological filters were applied to limit the retrieval to health technology assessments, systematic reviews, and meta-analyses, randomized controlled trials, and non-randomized studies. The search was also limited to English language documents published between January 1, 2015 and March 12, 2018.

Rapid response reports are organized so that the evidence for each research question is presented separately.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text publications was based on the selection criteria presented in Table 1.

Table 1: Selection Criteria

Population	Adult patients with a primary diagnosis of: -Major depressive disorder -Post-traumatic stress disorder -Generalized anxiety disorder
Intervention	Assessment (Q1) and psychological treatment (Q2) delivered via: -Telephone -Videochat (e.g., Skype, FaceTime, or other secure videoconferencing methods)
Comparators	Traditional in-person psychological assessment and treatment Wait list Treatment as usual Before/after treatment
Outcomes	Q1 - adequate assessment and treatment planning Q2 - improvements in mental health/condition, increased functioning, return to work, success in processing trauma
Study Designs	Health technology assessments, systematic reviews, meta-analyses, and randomized controlled trials

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, if they were duplicate publications, or if they were published prior to 2015. Additionally, an RCT was not eligible for our review if it had been included in one of the included SRs.

Studies that did not report on direct patient and provider interaction in real-time were not eligible for inclusion. Studies reporting the results of on-line or computer-based therapies, smart phone applications, text messaging, or automated self-management interventions were not included.

Critical Appraisal of Individual Studies

The included SRs were critically appraised using the AMSTAR 2 tool[12] and the Downs and Black checklist was used for quality assessment of included RCTs.[13] Summary scores were not calculated for the included SRs and RCTs; instead, a review of the strengths and limitations of each included study were described narratively. Additional details are available in Appendix 3: Tables 4 and 5.

Summary of Evidence

Quantity of Research Available

A total of 370 citations were identified in the literature search. Following screening of titles and abstracts, 326 articles were excluded and 44 potentially relevant reports from the electronic search were retrieved for full-text review. One potentially relevant publication was retrieved from the grey literature search. Of these 45 potentially relevant articles, 31 publications were excluded for various reasons, while 14 publications met the selection criteria and were included in this report. These comprised four SRs [10,14-16] and nine RCTs. [17-26] Two publications[21,22] were included for one study because the individual publications reported different outcomes. Appendix 1 presents the PRISMA flowchart of the study selection.

Summary of Study Characteristics

The body of evidence included four SRs,[10,14-16] of which two also included meta-analyses,[14,16] and nine RCTs (reported in ten citations),[17-26] that addressed the clinical effectiveness of telepsychotherapy in patients with depression,[14,17-23] anxiety,[24] depression and anxiety,[15,25] and PTSD.[10,16,26] Study characteristics are summarized below. Additional details regarding the characteristics of the included publications are available in Appendix 2: Table 2 and Table 3.

Study Design

Systematic Reviews

In all of the SRs,[10,14-16] comprehensive searches in two or more electronic bibliographic databases were performed. The date ranges covered by the searches were explicitly stated in two SRs: January 1991 to May 2016[15] and no restriction to publication date.[10] The other two SRs only reported the timeframe of the literature searches: June 2011 to December 2013[14] and January 2014 to March 2014.[16] All of the SRs included eligibility criteria for participants, descriptions of the interventions, comparators, and outcomes, and provided explanations for the selection of study designs. One SR[14] had the protocol for the review published four years earlier.[27] One SR[14] included only RCTs, whereas the other three SRs included RCTs and quasi-experimental and uncontrolled trials,[15] experimental study designs and non-experimental study designs (e.g., qualitative methods, case studies, cost analyses, or descriptive studies),[10] and clinical trials with formal eligibility requirements and routine care trials.[16] Two of the SRs[14,16] used meta-analyses to synthesize data from the included studies. The quality of the included studies was assessed in three SRs using various tools such as the Effective Public Health Practice Project Quality Assessment Tool,[15] Quality Assessment Tool for Quantitative Studies,[10] and the Cochrane Collaboration Risk of Bias tool.[14] Quality of included studies and risk of bias were not assessed in one SR.[16] The numbers of patients included in the two SR/meta-analyses were 5,159[14] and

1,191,[16] while the number of patients enrolled in the included trials of the SRs that did not conduct meta-analyses ranged from 6 to 127[15] and from 1 to 600.[10]

Randomized Controlled Trials

All nine RCTs[17-26] were multicentre trials, five were assessor or interviewer-blinded,[18-22,24] three were open-label,[23,25,26] and one RCT[17] stated that the teletherapists and the investigator who conducted the randomization were blinded to pre-intervention data. One RCT[23] was a choice-stratified RCT in which after randomizing nine patients, the investigators instituted a *"choice or equipoise-stratified"* (that incorporated individual preferences into randomization) approach to enhance the *"accessioning and ecological validity of the study."* One RCT[26] was a pilot study. Two RCTs[18,21,22] were designed as non-inferiority trials with pre-specified non-inferiority margins. Sample sizes ranged from 18[26] to 609 patients,[19] and follow-up times ranged from six weeks[26] to 24 months.[24]

Year of Publication and Country of Origin

Systematic Reviews

Two of the SRs[10,15] were published by authors located in the United Kingdom, while the other two SRs were published by authors located in Germany[14] and in the United States.[16] The publication years ranged from 2015[14,16] to 2018.[15]

Randomized Controlled Trials

Seven of the RCTs were published by authors located in the United States,[17,18,20-24,26] whereas the other two RCTs were published by authors located in England[19] and in Germany.[25] The publication years ranged from 2015[21,23,25] to 2017.[17,24,26]

Patient Population

Systematic Reviews

SRs of adult patients (≥ 18 years of age) were included, in keeping with the selection criteria for this report. One SR[14] focused on studies of adult patients with unipolar depressive disorders recruited from primary care settings. Another SR[15] considered studies in which adult patients received an intervention to reduce symptoms of major depression and/or anxiety. Two SRs in PTSD included ex-service military veterans with PTSD[10] and veterans or service members with PTSD attributed to combat-related trauma.[16]

Randomized Controlled Trials

Six RCTs[17-23] enrolled patients with major depression, of which three RCTs[17,20-22] stated that the diagnosis of major depressive disorder was made based upon Diagnostic and Statistical Manual IV (DSM-IV) criteria. The other three RCTs stated the diagnosis was based on other instruments such as the Patient Health Questionnaire 9-item [PHQ-9][19,23], Structured Clinical Interview for the DSM-IV Axis I Disorders [SCID],[23] or was not explicitly stated.[18] Of the RCTs in major depression, two RCTs[18,21,22] included only veterans and/or military service members and one RCT[20] included only Chinese-American immigrants. Two RCTs included patients with a diagnosis of major depression and also AIDS or HIV-infection[17] or a past history of traumatic brain injury.[23]

One RCT[24] included adult patients with generalized anxiety disorder and one RCT[25] included adult patients with depression or anxiety disorder. One pilot RCT[26] included adult

veterans meeting DSM-IV criteria for PTSD from military-related trauma associated with insomnia.

Interventions and Comparators

Systematic Reviews

All of the SRs[10,14-16] included a similar range of interventions (e.g., CBT, interpersonal therapy [IPT], behavior activation therapy for depression [BATD], prolonged exposure therapy [PE]) that was delivered via telephone or videoconferencing technology. In one SR[14] of depression, the comparators were usual care, placebo, or no treatment whereas in another SR[15] in depression and anxiety, the comparators included usual care, no treatment, benchmark comparison group, face-to-face therapy, waitlist control, or delayed treatment. In the two SRs[10,16] in PTSD, the comparators of interest were the same intervention (e.g., CBT, BATD, or PE) delivered by in-person therapy.

Randomized Controlled Trials

In the six RCTs[17-23] in depression, the interventions comprised IPT,[17] BATD,[18,21,22] and CBT delivered weekly over 6 to 9 sessions.[23] In these studies the comparators were standard care[17] or the same intervention delivered in-person (i.e. in the same room or face-to-face)[18,21,22], or with both the same intervention in-person or usual care.[23] Two RCTs[19,20] investigated integrated or collaborative approaches to delivering telehealth. In one RCT[19] the intervention was regular telephone contact from non-clinical trained health advisers to assess and advise patients with depression and the comparator was usual care. In another RCT[20] the intervention was a telepsychiatry-based culturally sensitive collaborative treatment (T-CSCT) comprising an assessment and a treatment component delivered via videoconferencing which was compared with treatment as usual (TAU).

In the one RCT[24] pertaining to anxiety, the intervention was telephone-delivered stepped collaborative care, which was compared with usual care. In the RCT[25] of patients with depression and anxiety, there were two teletherapy interventions (i.e., regular proactive weekly telephone contacts and the same weekly telephone contacts plus text messaging) compared to usual care. In the RCT[26] in PTSD, the intervention was telephone delivered CBT compared with CBT delivered in-person.

Outcomes

Depression

Outcomes of interest in patients with major depression were changes from baseline or the comparison of pre- and post-treatment scores using validated depression scales such as the Beck Depression Inventory (BDI-II),[14,15,17,21-23] Beck Hopelessness Scale (BHS),[18] Hamilton Depression Rating Scale (HAM-D or HAM-D17),[14,15,20,23], PHQ-9,[15,19] and the SCID.[15,18,21-23] The BDI-II includes 21 items with four response categories for each, resulting in a summated score ranging from 0 to 63, with higher scores indicating higher levels of depression symptom severity. The BHS contains 20 true-false statements relating to feelings of hopelessness about the future; sum scores can range from 0 to 20, with higher scores indicating higher levels of hopelessness. The PHQ-9 has a range of 0 to 27, with scores ≥ 10 to 14 indicating moderate depression, 15 to 19 moderately severe depression, and 20 to 27 severe depression. The SCID is a semi-structured interview guide used by clinicians or trained mental health professionals to classify and diagnose patients

as having major depressive disorder according to DSM-IV classification and diagnostic criteria.

In one SR and meta-analysis,[14] standardized mean differences (SMDs) were calculated for post-treatment depression scores by subtracting the mean score in the control group from the mean score in the intervention group and dividing by the pooled standard deviation (SD) to allow comparisons between treatment groups. A negative SMD indicates a better outcome due to the intervention. In the same SR,[14] responder and remission analyses were conducted where pooled treatment effects were reported as odds ratios (ORs) with 95% confidence intervals (CIs) and number needed to treat (NNT). In one SR,[15] Cohen's *d* was used to examine treatment effect sizes.

In two included RCTs (reported in three citations)[18,21,22] non-inferiority of the intervention (BATD by videoconferencing or videophone) was assessed with BATD delivered in-person. The treatment effect was based on changes from baseline in the BDI-II and BHS[18] or the BDI, SCID, or Geriatric Depression Scale (GDS).[21,22]

Other instruments that were reported mainly as secondary outcomes included the Montgomery-Asberg Depression Rating Scale,[14] Inventory of Interpersonal Problems,[17] Provision of Social Relations Scale,[17] various health-related quality of life instruments (e.g., EuroQol 5-Dimensions 5-Levels [EQ-5D-5L], Short Form [36] Health Survey [SF-36], Quality of Life Enjoyment and Satisfaction Questionnaire [Q-LES-Q]),[19,20,23] and others. Pre-specified subgroup analyses based on treatment modality (e.g., face-to-face and remote therapist-led) were performed in one SR.[14]

Anxiety

Outcomes of interest in patients with anxiety were health-related quality of life as measured by the SF-36 mental component summary (MCS) score in one RCT,[24] as well as various anxiety scales such as the Hamilton Anxiety Rating Scale (SIGH-A), Panic Disorder Symptom Scale (PDSS), Generalized Anxiety Disorder Severity Scale (GADSS), and the PHQ-9.[24] In the one RCT[25] of depression and anxiety, the primary outcome was the change in severity of anxiety, depression, and somatization as measured by the symptom scales of the German Brief Symptom Inventory Short Version (BSI-18).

Post-Traumatic Stress Disorder

Outcomes of interest in patients with PTSD included reduction of PTSD symptoms as measured by various scales such as the clinician-administered PTSD Scale or the PTSD Checklist.[10] PTSD-insomnia was evaluated by the self-reported Pittsburg Sleep Quality Index (PSQI) in one pilot RCT.[26] The PTSD checklist is a self-reported measure that evaluates 17 DSM-IV PTSD symptoms across three primary symptom clusters using a 5-point Likert-type scale; a total score of 50 is typically the threshold for identifying probable PTSD in those with military-related trauma.[18]

Managing risk and safety (i.e., based on suicidality or general patient safety in the context of high emotional levels in trauma-focused therapies) and therapeutic alliance was examined in one SR.[10]

Dropout rate (and correlates or predictors of dropout) was the main outcome of interest in one SR in PTSD.[14]

Safety outcomes were not assessed in any of the four SRs,[10,14-16] or in three of the RCTs.[24-26]

Summary of Critical Appraisal

Detailed summaries regarding the strengths and limitations of the included SRs and RCTs are provided in Appendix 3: Table 4 and Table 5.

Systematic Reviews

In all four SRs,[10,14-16] an 'a priori' design was described, comprehensive literature searches of 3 to 5 databases were conducted, and lists and the characteristics of included studies were provided. None of the SRs provided lists of excluded studies in the publications. The quality of the included studies and risk of bias was assessed in all but one SR[16] and in general, approximately one third of included studies were considered to be strong, another third to be moderate or uncertain, and a third were considered weak. An explanation for the selection of study design was provided in all but one SR.[10] Follow-up periods for the included studies were not reported in any of the SRs. The procedures for study selection and data extraction by reviewers were not provided in one SR[10] and were unclear in another.[16] All but one SR[10] investigated and/or discussed study heterogeneity. Two SRs[14,16] incorporated meta-analyses and both used appropriate methodology to combine study data. One SR[10] reported all results narratively. One SR[16] did not specify all the comparators used in the included studies; however, comparator information was available for the four included studies in this SR that utilized telemedicine. None of the SRs reported on safety outcomes. Sources of funding for individual studies was not provided in any of the SRs. Conflict of interest was not reported by the authors in one SR.[16]

Randomized Controlled Trials

The study objectives, screening criteria, inclusion/exclusion criteria, interventions, comparators, patient characteristics, and main outcome measures were clearly described in all nine RCTs.[17-26] Appropriate measures of random variability and statistical tests to assess the main study outcomes were used in all the RCTs. All but one RCT[26] analyzed outcomes using intention-to-treat principles. Six RCTs[17-23] stated that adverse events were either not reported during the study, or were monitored or investigated, whereas three RCTs[24-26] did not report or mention any safety outcomes. Limitations identified for the included RCTs were the lack of blinding of patients and treatment providers,[18-26] although in five RCTs the outcome assessor or interviewer were blinded.[18-22,24] The generalizability of the study findings to other populations may be limited in the RCTs of patients with AIDS/HIV infection,[17] traumatic brain injury,[23] or that enrolled only Chinese-American immigrants. Similarly, the results of RCTs that included predominantly males (68% to 98%),[20-22] or only male veterans[26] which comprise more than a third of the included RCTs, may not be generalizable to females. Although patient disposition was reported in all the RCTs, the number of patients lost to follow-up was not reported in two RCTs.[17,20] and was unclear in two RCTs.[23,24] In one RCT[21,22] the authors commented that the technology used in the study was considered somewhat obsolete at the time of study completion.

Summary of Findings

The overall findings of this review are summarized below. Additional details are available in Appendix 4: Table 6 and Table 7.

1. *What is the clinical effectiveness of psychological assessment delivered via telehealth in patients with depression, PTSD, or anxiety?*

None of the included SRs addressed the psychological assessment of patients with depression, anxiety, or PTSD, as they all focused on treatment. Of the included RCTs, two studies[19,20] that investigated integrated or collaborative telehealth interventions for depression incorporated a patient assessment component as part of the collaborative telehealth intervention; however, the measured outcomes encompass both the effects of assessment and treatment, so the findings are considered to be inconclusive.

In one RCT,[19] patients were randomly assigned to a teleheath service plus usual care or usual care alone. The intervention comprised regular telephone calls to patients from non-clinical, but trained, health advisers who contacted patients after an initial assessment on six separate occasions over four months and followed up with reinforcement or to monitor for relapse for an additional two months. Using scripts and interactive software, the advisers supported and directed patients based on their assessment of patient need to on-line resources, interactive programs including computerized CBT therapy and other information sources and tools. The primary outcome was the proportion of patients with response defined as a PHQ-9 score ≤ 10 and a reduction of ≥ 5 points at four weeks. At four months, more patients (27%) in the intervention group had a treatment response compared with 19% of those in the usual care group, which was statistically significantly different. In comparison with usual care alone, patients reported statistically significant improvement in measures of anxiety, treatment satisfaction, access to care, and satisfaction with the support received.

The other RCT[20] investigated T-CSCT compared with TAU to provide culturally sensitive telepsychiatry-based assessment and psychiatric consultation to monolingual Chinese-American immigrants with major depressive disorder over a six month period. As part of the collaborative care model, patients were assigned a clinical care manager who monitored and assessed the patient's depressive symptoms, adherence to their treatment protocol, and patient's self-management of their depression. Patients in the TAU group received an initial telepsychiatry-based assessment with treatment recommendations but did not receive culturally-sensitive consultation or on-going monitoring and care management. The primary outcome variables were response and remission rates based on the HAM-D17 score at the last assessment compared to baseline. Response was defined as ≥ 50% improvement in the HAM-D17 score and remission was defined as having a HAM-D17 score ≤ 7 at the last measurement. A statistically significantly larger proportion of patients in the T-CSCT intervention group had response (45.0%) and remission (32.2%) compared to TAU (17.0% and 10.4%, respectively; $P < 0.001$ for both). Secondary outcomes based on clinical global impression scales also significantly favoured T-CSCT; however, there was no statistical difference in Q-LES-Q scores between the treatment groups.

No relevant evidence was identified for psychological assessment delivered via telehealth for patients with anxiety or PTSD.

2. *What is the clinical effectiveness of psychological treatment delivered via telehealth in patients with depression, PTSD, or anxiety?*

Depression

One SR and meta-analysis[14] and six RCTs[17-23] investigated psychotherapy delivered via telehealth for the treatment of major depression. One SR[15] and one RCT[25] studied telehealth interventions in patients with both depression and anxiety.

Post-treatment depression scores

Based on the results of the SR and meta-analysis,[14] both face-to-face CBT and remote therapist-led CBT were found to have a statistically significant treatment effect (better

outcomes) when compared with control groups, but the difference between the two types of delivering therapy was not statistically significant. A pre-specified subgroup analysis by delivery mode revealed no difference in treatment effect when therapy was delivered face-to-face (-0.23 [-0.33 to -0.14] or for remote therapist-led treatment (-0.43 [-0.62 to -0.25]). Differences between the effects of all the different interventions included in the meta-analysis were not statistically significant (*P*=0.74), nor was the test for subgroup differences statistically significant (*P*=0.14).

In the SR[15] in both depression and anxiety, results were compared using Cohen's *d*, which is an effect size used to indicate the standardized difference between two means. For 5 of the 6 included RCTs of CBT or IPT by telephone compared with no treatment, enhanced usual care, or TAU, the treatment effect was statistically significant and Cohen's *d* ranged from 0.25 to 1.98 (median 1.58). Similarly the included quasi-experimental and uncontrolled studies also showed statistically significant reductions in depression scores after telephone-delivered CBT.

Two RCTs[17,23] studied telehealth in patients diagnosed with major depression, but in specialized populations. In one RCT,[17] telephone-administered IPT was investigated to determine if the treatment acutely relieved depressive symptoms in AIDS/HIV-infected patients with major depressive disorder. A series of ITT, completer, and sensitivity analyses found that across all analyses, patients who received telephone-IPT had significantly less depressive symptoms as measured by the BDI-II compared with standard care. Similar results were found for interpersonal problems, as measured with the Inventory of Personal Problems scale. In a RCT[23] of patients with major depression and history of traumatic brain injury, various instruments (HAM-D 17, SCL-20, Physician Global Impression) demonstrated comparable results between CBT-T, CBT-IP or usual care.

One RCT[19] compared treatment response (as measured by the PHQ-9 at four months) between a telephone intervention plus usual care versus usual care alone. Treatment response was higher for the intervention plus usual care (27%) compared to usual care alone (19%), which was statistically significantly different.

Response and remission

In the SR and meta-analysis,[14] ORs were calculated for the dichotomous outcomes of response and remission (defined In Appendix 4, Table 6). An OR > 1 indicates more events in the intervention group compared with the control group. For both face-to-face CBT and remote therapist-led CBT the ORs for response and remission all exceeded 1, thus indicating more events in the intervention groups and implying similar treatment effects between face-to-face and telehealth interventions.

In a RCT[20] of a collaborative treatment approach (T-CSCT) delivered by videoconferencing to Chinese American immigrants, both response (45.0%) and remission (32.3%) were statistically significantly higher in patients who received T-CSCT compared with TAU (17.0% and 10.4%, respectively; *P* < 0.001 for both).

Non-inferiority

Two RCTs[18,21,22] assessed the non-inferiority of BATD delivered via videoconferencing/videophone compared to in-office/face-to-face treatment based on post-treatment BHS and BDI-II scores[18] or BDI and GDS scores[21,22] using both ITT and per-protocol analyses. Demonstration of non-inferiority means that the intervention is not worse than the comparator by more than a small pre-specified amount. In both studies, non-

inferiority of the telehealth treatment to in-person treatment was demonstrated. No statistically significant differences were found between the different treatment delivery modes for numerous secondary outcomes measured by other scales,[18,21,22] patient satisfaction[18,21,22] or quality of life as measured by the SF-36.[21,22]

Treatment discontinuation

Pooled estimates for study discontinuation as a dichotomous outcome were also examined in the SR and meta-analysis.[14] For both face-to-face CBT and remote therapist-led CBT, the ORs were very close to 1 (implying a similar number of events in the intervention group compared with the control group) and no statistically significant differences between the two types of therapy delivery. In the RCTs[17-19,21-23] where information on patient disposition post-randomization was available, discontinuation rates ranged from 14%[19] to 32.3%, and appeared to be similar in the majority of studies.[18].

Anxiety

Post-treatment anxiety scores

In the SR[15] that included both depression and anxiety, three RCTs were included that reported statistically significant reductions in anxiety symptoms following telephone-delivered interventions. Effect sizes in two of the studies based on Cohen's *d* ranged from 0.34 to 1.07 (median 0.69). One quasi-experimental study included in the SR compared telephone-delivered CBT with a delayed wait-list control group and found significant reductions in symptoms compared with the control group.

In a RCT[25] of patients with both depression and anxiety, pro-active regular telephone contact combined with text messaging significantly reduced anxiety scores after six months, as measured by the BSI-18 scale ($P = 0.042$). In a sensitivity analysis restricted to 75% of the patients with the highest symptom scores at baseline, telephone contact alone (without text messaging) also had a significant effect for both anxiety ($P = 0.036$) and depression ($P = 0.046$)

Health-related quality of life

In the single RCT[24] examining interventions for anxiety, at 12 months, a telephone-delivered stepped collaborative care intervention resulted in statistically significant improvements compared to usual care in quality of life as measured by the SF-36 mental component summary, as well as various anxiety scales (SIGH-A, PDSS, GADSS, and PHQ-9). Effect sizes were diminished at 24 months.

PTSD

Two SRs[10,16] and one RCT[26] investigated trauma-focused therapies delivered by telehealth compared with in-person treatment of the same intervention.

Post-treatment PTSD symptom scores

One SR[10] reported results narratively; however, with regard to post-treatment outcomes, 15 of 18 included studies that evaluated teletherapy versus a control (typically the same intervention but in-person) found that teletherapy was associated with significant reductions in PTSD symptoms, regardless of the type of psychotherapy intervention, with the exception of one study that measured only anger. In 12 of the 18 studies, non-inferiority was assessed, and of these nine studies it was concluded that teletherapy was as effective

as in-person therapy. These results were supported by none of the included studies reporting significant differences in patient satisfaction and acceptability between the delivery modes.

In a pilot RCT[26] of PTSD symptom-related insomnia, telephone-delivered CBT was shown to decrease PTSD-related insomnia symptoms over a three month period in a similar manner as in-person CBT therapy, as measured by the PSQI. The small sample size (n=18) precluded comparisons being made between treatment groups.

Management of risk and safety

In the SR[10] that reported the results narratively, four included studies reported on issues relating to suicidality or general patient safety in the context of high emotional levels arising in trauma-focused therapies. Of these, three studies reported no significant difficulties in using teletherapy to manage the situation whereas one study, while not reporting significant differences, did report a trend for participant concern about managing emotions during exposure tasks via teletherapy (i.e., without the physical presence of the therapist).

Treatment discontinuation

One SR[16] focused solely on dropout rates and correlates or predictors of dropout from clinical trials or routine care trials of interventions for PTSD. The overall pooled dropout rate (95% CI) based on 20 included studies was 36.0% (26.90 to 45.00). When dropout rate was examined by treatment modality (i.e., telemedicine versus in-person treatment) the OR was very close to 1 and the difference between modalities was not statistically significant. Identified correlates of dropout were age (younger), being unmarried, unemployed, having service-connected disability, and greater PTSD severity. In the other SR,[10] in 13 included studies, no significant differences in attrition between teletherapy and in-person treatments were found in any of the included studies.

Limitations

There are various limitations associated with the evidence in our report for the use of telehealth to deliver psychotherapy. A key limitation is the diversity and large number of scales and instruments that were used to measure treatment effects in the included SRs and RCTs. The wide range of treatment outcomes, methods of data presentation and analysis, coupled with differences in the statistical methodology, made comparisons between studies difficult. Furthermore, many of the scales and instruments used were self-administered whereas many others were completed by clinicians or assessors/interviewers and therefore subject to different inherent biases. The types of study designs also varied considerably, as evidenced by the studies included in the SRs, as only one SR[14] included RCTs whereas the other three SRs[10,15,16] included a mixture of RCTs and experimental or quasi- or non-experimental study designs. The population of interest was also broad, including adult patients with major depression, generalized anxiety disorder, and PTSD, conditions that can exist separately, or frequently overlap with each other, or are associated with other comorbidities. All of the included RCTs are also compromised by patients who were not blinded to the treatment allocation, which it is acknowledged may not be possible due to the nature of the intervention (telephone or videoconferencing) and comparators (in-person/same room or usual care). In five of the included RCTs,[18-22,24] assessors were blinded to the treatment allocation in attempts to address detection bias.

There was limited evidence available for the first research question on the clinical effectiveness of psychological assessment delivered via telehealth. Two RCTs[19,20] that

investigated collaborative care approaches delivered by telehealth included an assessment component; however, the measured outcomes are based on the combined effect of the assessment and treatment and so the reported findings are not specific to psychological assessment alone. No relevant evidence was identified for psychological assessment delivered via telehealth for patients with anxiety or PTSD.

Findings from the included RCTs that enrolled patients with major depression and AIDS/HIV-infection,[17] traumatic brain injury,[23] or only Chinese-American immigrants[20] may have limited generalizability to other populations. None of the studies included in this report were conducted in Canada, with the exception of the inclusion of one Canadian study in a SR[10] on PTSD. Therefore, it is not clear to what extent the results would be applicable to a Canadian setting, although the majority of RCTs[17,18,20-24,26] and one SR[16] were authored by researchers located in the United States, and so may present somewhat of a North American perspective.

Lastly, due to the rapidly evolving field of telecommunications, it must be considered that the technology for a telehealth intervention that was included in this report, may no longer be current, and in fact may be obsolete, as acknowledged in one included study,[21,22] and thus no longer applicable to current clinical practice.

Conclusions and Implications for Decision or Policy Making

The current report, which was intended to be an update of a previous CADTH Rapid Response review,[11] summarizes the results of four SRs[10,15] (two of which included meta-analyses)[14,16] and nine RCTs.[17-26] The previous report did not summarize the evidence.

Limited evidence suggests that psychological assessment delivered via telehealth is clinically effective in patients with depression. Two RCTs[19,20] that investigated collaborative care approaches delivered by telehealth included an assessment component; however, the measured outcomes are based on the combined effect of the assessment and treatment and so the findings are inconclusive. No relevant evidence was identified for psychological assessment delivered via telehealth for patients with anxiety or PTSD.

Despite the stated limitations, there was concurrence among the included SRs and RCTs that the evidence supports that psychological treatment delivered by telehealth (telephone or videoconferencing/videophone) is clinically effective and that the magnitude of the treatment effect is comparable between telehealth and the same treatment delivered by traditional means (in-person/same room therapy) for adult patients with depression, anxiety, or PTSD. Two included RCTs[18,21,22] also reported that psychotherapy delivered via videoconferencing/videophone was non-inferior to the same therapy provided in-office or face-to-face, based on post-treatment depression scores. Limited evidence from two SRs in depression[14] and PTSD[16] supports that discontinuation rates do not differ between patients who receive treatment via telehealth compared to traditional in-person therapy.

Further research is required to confirm the clinical effectiveness of telehealth for the psychological assessment of patients with depression, anxiety, or PTSD. Another consideration is that due to rapidly evolving developments in telecommunication technology, the evidence to support the clinical effectiveness of a particular technology may be outdated or perhaps not yet available, both of which have implications for decision and policy making.

References

1. Totten AM, Womack DM, Eden KB, McDonagh MS, Griffin JC, Grusing S, et al. Telehealth: mapping the evidence for patient outcomes from systematic reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2016 Jun. (AHRQ Comparative Effectiveness Technical Briefs).

2. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open. 2017 Aug 3;7(8):e016242. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5629741

3. Cunningham JEA, Shapiro CM. Cognitive behavioural therapy for insomnia (CBT-I) to treat depression: a systematic review. J Psychosom Res. 2018 Mar;106:1-12.

4. Engel CC, Litz B, Magruder KM, Harper E, Gore K, Stein N, et al. Delivery of self training and education for stressful situations (DESTRESS-PC): a randomized trial of nurse assisted online self-management for PTSD in primary care. Gen Hosp Psychiatry [Internet]. 2015 [cited 2018 Mar 15];37(4):323-8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4762212 /

5. Brabyn S, Araya R, Barkham M, Bower P, Cooper C, Duarte A, et al. The second Randomised Evaluation of the Effectiveness, cost-effectiveness and Acceptability of Computerised Therapy (REEACT-2) trial: does the provision of telephone support enhance the effectiveness of computer-delivered cognitive behaviour therapy? A randomised controlled trial. Health Technol Assess [Internet]. 2016 Nov [cited 2018 Mar 15];20(89):1-64. Available from: https://www.journalslibrary.nihr.ac.uk/hta/hta20890/#/abstract

6. Canadian Medical Association, Canadian Psychiatric Association. Joint statement on access to mental health care [Internet]. Ottawa: Canadian Medical Association; 2016. [cited 2018 Mar 22]. Available from: https://www.cma.ca/Assets/assets-library/document/en/advocacy/policy-research/cma-policy-access-to-mental-health-care-pd16-04-e.pdf

7. CANSIM [Internet]. Ottawa: Statistics Canada; 2013 Sep 18. Table 105-1101 mental health profile, Canadian Community Health Survey - Mental Health (CCHS), by age group and sex, Canada and provinces; 2013 [cited 2018 Mar 22]. Available from: http://www5.statcan.gc.ca/cansim/a26?lang=eng&id=1051101

8. Canadian Forces mental health survey, 2013 [Internet]. Ottawa: Statistics Canada; 2014 Aug 11 [cited 2018 Mar 22]. Available from: http://www.statcan.gc.ca/daily-quotidien/140811/dq140811a-eng.pdf

9. Marton K, Kanas N. Telehealth modalities for group therapy: Comparisons to in-person group therapy. Int J Group Psychother. 2016;66(1):145-50.

10. Turgoose D, Ashwick R, Murphy D. Systematic review of lessons learned from delivering tele-therapy to veterans with post-traumatic stress disorder. J Telemed Telecare. 2017 Jan 1;1357633X17730443.

11. Tele-medicine for patients with mental health disorders: clinical and cost-effectiveness [Internet]. Ottawa (ON): CADTH; 2015 [cited 2018 Mar 13]. Available from: https://www.cadth.ca/sites/default/files/pdf/htis/aug-2015/RA0796%20Telemedicine%20for%20mental%20health%20Final.pdf

12. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ [Internet]. 2017 [cited 2018 Mar 15];358:j4008. Available from: http://www.bmj.com/content/bmj/358/bmj.j4008.full.pdf

13. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health [Internet]. 1998 Jun [cited 2018 Mar 15];52(6):377-84. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1756728/pdf/v052p00377.pdf

14. Linde K, Sigterman K, Kriston L, Rucker G, Jamil S, Meissner K, et al. Effectiveness of psychological treatments for depressive disorders in primary care: systematic review and meta-analysis. Ann Fam Med [Internet]. 2015 Jan [cited 2018 Mar 15];13(1):56-68. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4291267/

15. Coughtrey AE, Pistrang N. The effectiveness of telephone-delivered psychological therapies for depression and anxiety: a systematic review. J Telemed Telecare. 2018 Feb;24(2):65-74.

16. Goetter EM, Bui E, Ojserkis RA, Zakarian RJ, Brendel RW, Simon NM. A systematic review of dropout from psychotherapy for posttraumatic stress disorder among Iraq and Afghanistan combat veterans. J Trauma Stress. 2015 Oct;28(5):401-9.

17. Heckman TG, Heckman BD, Anderson T, Lovejoy TI, Markowitz JC, Shen Y, et al. Tele-interpersonal psychotherapy acutely reduces depressive symptoms in depressed HIV-infected rural persons: a randomized clinical trial. Behav Med. 2017 Oct;43(4):285-95.

18. Luxton DD, Pruitt LD, Wagner A, Smolenski DJ, Jenkins-Guarnieri MA, Gahm G. Home-based telebehavioral health for U.S. military personnel and veterans with depression: a randomized controlled trial. J Consult Clin Psychol. 2016 Nov;84(11):923-34.

19. Salisbury C, O'Cathain A, Edwards L, Thomas C, Gaunt D, Hollinghurst S, et al. Effectiveness of an integrated telehealth service for patients with depression: a pragmatic randomised controlled trial of a complex intervention. Lancet Psychiatry. 2016 Jun;3(6):515-25.

20. Yeung A, Martinson MA, Baer L, Chen J, Clain A, Williams A, et al. The Effectiveness of telepsychiatry-based culturally sensitive collaborative treatment for depressed Chinese American immigrants: a randomized controlled trial. J Clin Psychiatry. 2016 Aug;77(8):e996-e1002.

21. Egede LE, Acierno R, Knapp RG, Lejuez C, Hernandez-Tejada M, Payne EH, et al. Psychotherapy for depression in older veterans via telemedicine: a randomised, open-label, non-inferiority trial. Lancet Psychiatry. 2015 Aug;2(8):693-701.

22. Egede LE, Acierno R, Knapp RG, Walker RJ, Payne EH, Frueh BC. Psychotherapy for depression in older veterans via telemedicine: effect on quality of life, satisfaction, treatment credibility, and service delivery perception. J Clin Psychiatry. 2016 Dec;77(12):1704-11.

23. Fann JR, Bombardier CH, Vannoy S, Dyer J, Ludman E, Dikmen S, et al. Telephone and in-person cognitive behavioral therapy for major depression after traumatic brain injury: a randomized controlled trial. J Neurotrauma [Internet]. 2015 Jan 1 [cited 2018 Mar 15];32(1):45-57. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4273196/

24. Rollman BL, Belnap BH, Mazumdar S, Abebe KZ, Karp JF, Lenze EJ, et al. Telephone-delivered stepped collaborative care for treating anxiety in primary care: a randomized controlled trial. J Gen Intern Med [Internet]. 2017 Mar [cited 2018 Mar 15];32(3):245-55. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5330997/

25. van den Berg N, Grabe HJ, Baumeister SE, Freyberger HJ, Hoffmann W. A Telephone- and text message-based telemedicine concept for patients with mental health disorders: results of a randomized controlled trial. Psychother Psychosom. 2015 Feb 21;84(2):82-9.

26. Franklin CL, Walton JL, Raines AM, Chambliss JL, Corrigan SA, Cuccurullo LJ, et al. Pilot study comparing telephone to in-person delivery of cognitive-behavioural therapy for trauma-related insomnia for rural veterans. J Telemed Telecare. 2017 Jan 1;1357633X17732366.

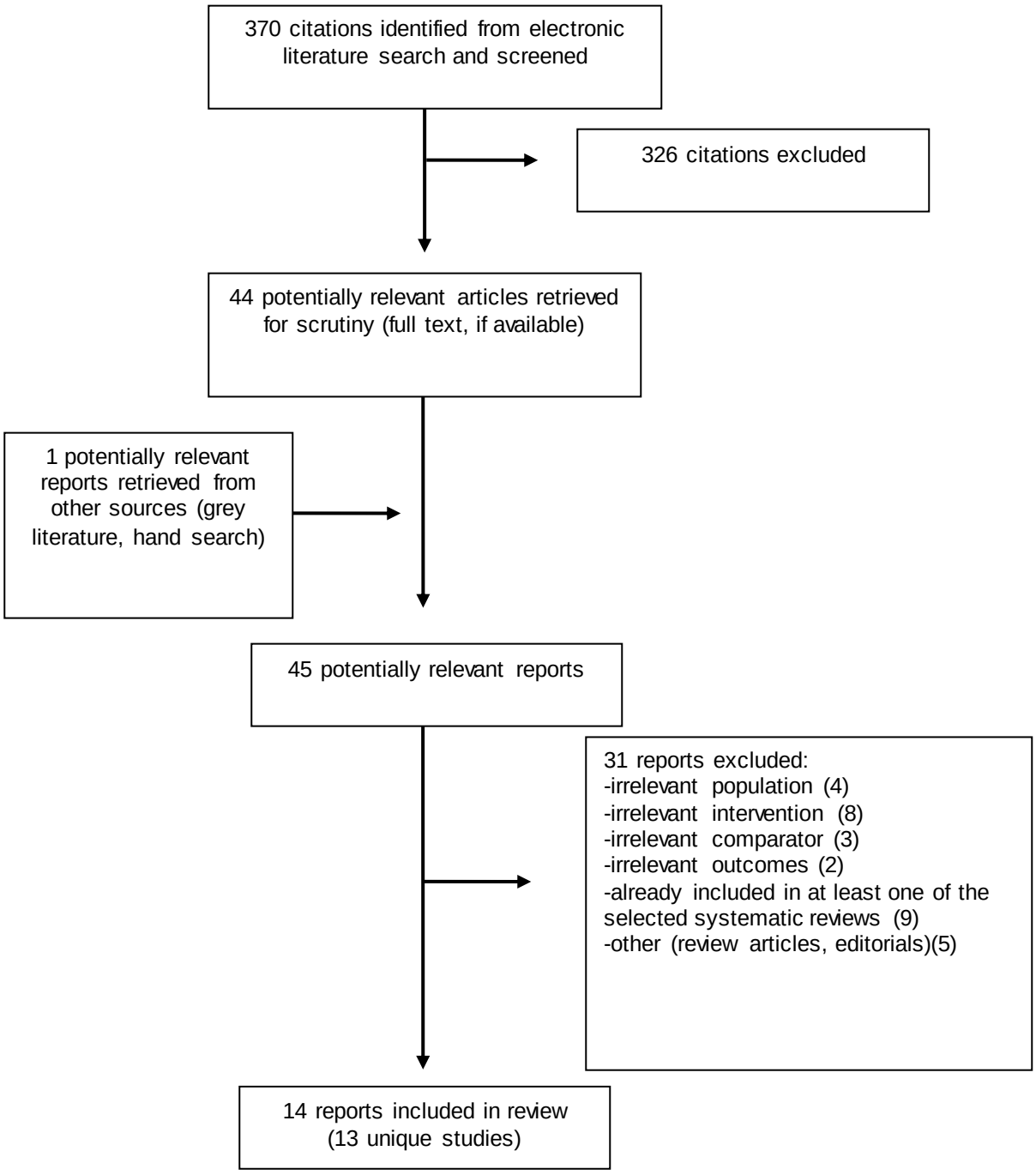
27. Linde K, Schumann I, Meissner K, Jamil S, Kriston L, Rucker G, et al. Treatment of depressive disorders in primary care--protocol of a multiple treatment systematic review of randomized controlled trials. BMC Fam Pract [Internet]. 2011 Nov [cited 2018 Mar 15];15;12:127. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3226438

28. Luxton DD, Pruitt LD, O'Brien K, Stanfill K, Jenkins-Guarnieri MA, Johnson K, et al. Design and methodology of a randomized clinical trial of home-based telemental health treatment for U.S. military personnel and veterans with depression. Contemp Clin Trials. 2014 May;38(1):134-44.

29. Thomas CL, Man MS, O'Cathain A, Hollinghurst S, Large S, Edwards L, et al. Effectiveness and cost-effectiveness of a telehealth intervention to support the management of long-term conditions: study protocol for two linked randomized controlled trials. Trials [Internet]. 2014 Jan 24 [cited 2018 Mar 26];15:36. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3906859

30. van den Berg N, Grabe HJ, Freyberger HJ, Hoffmann W. A telephone- and text-message based telemedical care concept for patients with mental health disorders--study protocol for a randomized, controlled study design. BMC Psychiatry [Internet]. 2011 Feb 17 [cited 2018 Mar 19];11:30. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3045884

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table 2: Characteristics of Included Systematic Reviews and Meta-Analyses

First Author, Year, Country	Study Design and Numbers of Primary Studies Included	Population Characteristics	Intervention	Comparator(s)	Reported Outcomes
Depression					
Linde, 2015[14] Germany	A systematic review and meta-analysis of 30 RCTs published from 1984 to 2013. The countries of origin of the studies were not provided.	5,159 adult (≥ 18 years old) primary care patients with unipolar depressive disorders	Psychological treatment for depression (4 included studies reported on remote-therapist led CBT or problem-solving therapy delivered by telephone)	Usual care, PBO, or no treatment	Efficacy -Effect measured as either change from baseline or post-treatment scores on a depression scale (e.g., BDI, HAM-D, Montgomery-Asberg Depression Rating Scale) -Response -Remission Other -Study discontinuation -Subgroup analysis based on treatment modality
Depression and Anxiety					
Coughtrey, 2018[15] United Kingdom	A systematic review of 14 studies (9 RCTs, 2 quasi-experimental design, and 3 uncontrolled designs). Studies were published from 1995 to 2014. The countries of origin of the included studies were not provided.	Adult (≥ 18 years old) patients who had received an intervention to reduce symptoms of depression and/or anxiety. Study sample sizes ranged from n=6 to n=127. (10 studies were in depression and 4 were in anxiety disorders)	Evidence-based (NICE-recommended) psychological therapy for depression and/or anxiety disorders delivered by telephone (i.e., CBT, IPT and BA for depression and CBT, ERP and applied relaxation for anxiety disorders)	Usual care, no treatment, benchmark comparison group, face-to-face therapy, waitlist control, or delayed treatment	Efficacy: -effect sizes as measured by Cohen's d based on self-reported questionnaire measures or clinician-assessed/rated outcomes (e.g., HAM-D, BDI-II, PHQ-9, SCID, SCL, QIDS-SR, POMS-DD, YBOCS)
Post-Traumatic Stress Disorder					
Turgoose, 2017[10] United Kingdom	A systematic review of 41 studies (28 studies had experimental study designs [not specified] and 13 studies had non-experimental study designs [qualitative methods, case	Ex-military personnel with PTSD receiving a teletherapy intervention. Study sample sizes ranged from n=1 to n=600.	Trauma-focused therapies (e.g., PE, BA, CBT) delivered by teletherapy (videoconferencing or telephone) (30 studies reported on the effectiveness of	Same intervention delivered in-person	Efficacy -Clinical effectiveness as measured by reduction in PTSD symptoms -Therapeutic alliance Safety -Managing risk (suicidality, general patient safety) Other

First Author, Year, Country	Study Design and Numbers of Primary Studies Included	Population Characteristics	Intervention	Comparator(s)	Reported Outcomes
	studies, cost analyses, or were descriptive]). Studies were published from 2004 to 2017. The country of origin was the USA for all but one Canadian study.		teletherapy interventions directly compared to in-person interventions)		-Drop-out and attendance rates -Patient satisfaction and acceptability -Technology use
Goetter, 2015[16] United States	A systematic review and meta-analysis of 20 studies (10 studies were described as clinical trials with formal eligibility requirements and 10 were trials of routine clinical care). Studies were published from 2009 to 2014. The countries of origin of the included studies were not provided; however, the inclusion criteria required all participants to be OEF/OIF/OND veterans or service members which are US-led operations.	1,191 veterans or service members with PTSD for combat-related trauma from OEF/OIF/OND receiving a psychotherapy intervention in an outpatient setting	Psychotherapy treatment (i.e., PE, CBT, CPT) in an outpatient setting (4 of the included studies reported on the use of telemedicine to deliver BA or PE therapy)	Not specified for all included studies (In the 4 studies of telemedicine, the comparators were the same intervention delivered in-person)	-Dropout rate -Correlates or predictors of dropout

BA = behavioral activation; BDI-II = Beck Depression Inventory; CAPS = Clinician Administered PTSD Scale; CBT = cognitive behavioral therapy; CPT = cognitive processing therapy; ERP = exposure and response prevention; HAM-D = Hamilton Depression Rating Scale; IPT = interpersonal psychotherapy; NICE = National Institute for Health and Care Excellence; OEF = US Operation Enduring Friendship; OIF = US Operation Iraqi Freedom; OND = US Operation New Dawn; PBO = placebo; PCL = PTSD Checklist; PE = prolonged exposure; PHQ-9 = Patient Health Questionnaire-9; POMS-DD = Profile of Mood States Depression-Dejection Scale; PTSD = Post-Traumatic Stress Disorder; QIDS-SR = Quick Inventory of Depressive Symptomatology; RCT = randomized controlled trial; SCID = Structured Clinical Interview; SCL = Hopkins Symptom Checklist; US = United States; Y-BOCS = Yale Brown Obsessive Compulsive Scale - Self Report Version

Table 3: Characteristics of Included Randomized Controlled Trials

First Author, Year, Country	Study Design (Follow-up)	Study Population	Intervention (Sample size)	Comparator(s) (Sample size)	Reported Outcomes
Depression					
Heckman, 2017[17] United States	Prospective, multicentre RCT (teletherapists and the investigator who conducted the randomization were blinded to patients' eligibility screening and pre-intervention data) (Surveys were completed at pre-intervention, post-intervention, and 4- and 8-month follow-up)	Inclusion criteria: Adult (≥ 18 years old) with self-reported AIDS or HIV-infection and a DSM-IV diagnosis of MDD, MDD in partial remission, or dysthymic disorder based on the mood module of the PRIME-MD, and living in a rural county for ≥ 1 year Exclusion criteria: serious cognitive or neurosychiatric impairment	T-IPT plus standard care (n=70) (9 weekly one hour T- IPT treatments delivered by a therapist)	Standard care (n=62)	Primary: Change between pre- and post-treatment depressive symptoms measured by the BDI-II Secondary: Change between pre- and post-treatment interpersonal problems and social support as measured by the IIP and PSRS, respectively -Responder analysis (This publication reported only acute changes from pre- to post-intervention)

First Author, Year, Country	Study Design (Follow-up)	Study Population	Intervention (Sample size)	Comparator(s) (Sample size)	Reported Outcomes
Luxton, 2016[18] United States	Prospective, multicentre, assessor-blinded, non-inferiority RCT (12 weeks)	Inclusion criteria: US military service members and veterans 18 to 65 years old meeting diagnostic criteria for minor or major depressive disorder, high-speed home Internet access (384 kBs/min), if taking psychoactive medication must be stable x 30 days Exclusion criteria: undergoing psychotherapy for depression, active psychotic symptoms/disorder, dysthymic disorder, suicidal ideation, organic mental disorder, substance dependence, history of violence or poor impulse control, significant on-going stressors requiring urgent crisis intervention, private space in home	BATD via videoconferencing in the home setting x 8 sessions (n=62)	BATD in a traditional in-office (same room) setting x 8 sessions (n= 59)	Primary: Non-inferiority of change from baseline in BHS and BDI-II to post-treatment at 8 weeks Secondary: -SCID-I/P -BAI -PCL-M -IASMHS -CSQ Safety: as recorded on a treatment session checklist: suicidal ideation, homicidal ideation, presence of firearm, signs of intoxication, disorientation, severe emotional dysregulation
Salisbury, 2016[19] England	Prospective, multicentre, pragmatic, assessor-masked RCT (12 months)	Inclusion criteria: Adults (≥ 18 years) with PHQ-9 score ≥ 10 and confirmed diagnosis of depression on CIS-R scale with access to a telephone, the Internet, and email Exclusion criteria: currently receiving therapy or case management, given birth ≤ 12 months, history of major bipolar disorder, psychotic illness, dementia, severe learning disability, substance dependency, palliative care, or significant suicide risk	Telehealth (telephone) service plus usual care x 6 sessions over approximately 4 months + up to 3 calls at 2 month intervals to provide reinforcement and to detect relapse (n=307)	Usual care (n=302)	Primary: Responders defined as the proportion of participants with PHQ-9 score ≤ 10 and reduction of ≥ 5 points) at 4 weeks Secondary: -PHQ-9 at 8 and 12 months -GAD-7 -EQ-5D-5L -Patient satisfaction -Access to care -Hei-Q -Adherence -eHEALs
Yeung, 2016[20] United States	Prospective, multicentre, blinded-	Inclusion criteria: Adult (≥ 18 years) monolingual Chinese-American immigrants meeting	T-CSCT via videoconferencing	TAU (n=93)	Primary: -Responders defined as a ≥

First Author, Year, Country	Study Design (Follow-up)	Study Population	Intervention (Sample size)	Comparator(s) (Sample size)	Reported Outcomes
	interviewer RCT (6 months)	DSM-IV criteria for MDD as diagnosed by the MINI, score of ≥ 10 on the CB-PHQ-9 Exclusion criteria: serious suicidal risk, unstable medical illness with a significant likelihood of requiring hospitalization during the study period, comorbid severe mental disorders, been treated by a psychiatrist in the past 4 months	(n=97) The T-CSCT involved 2 major components: culturally-sensitive psychiatric assessment (using the *Engagement Interview Protocol for Cultural Consultation*) and collaborative care		50% response on the HAMD17 score -Remission was defined as a HAMD17 score ≤ 7 Secondary: -CGI-S -CGI-I -Q-LES-Q
Egede, 2015,[21] 2016[22] United States	Prospective, multicentre, assessor-blinded, non-inferiority RCT (12 months)	Inclusion criteria: Veterans (≥ 58 years) meeting DSM-IV criteria for major depressive disorder Exclusion criteria: Actively psychotic or demented individuals, suicidal ideation and clear intent, substance dependence	BATD via telemedicine (videophone) x 8 weeks (n=120)	BATD via same-room (face-to-face) treatment (n=121)	Primary: Treatment response by GDS and BDI and SCID in the PP population at 12 months Secondary: Continuous BDI and GDS scores and response status at 4 and 8 weeks and 3 months
Fann, 2015[23] United States	Prospective, multicentre, choice-stratified, open-label, 3-arm RCT Randomization stratified on TBI severity and randomization option choice (16 weeks)	Inclusion criteria: Adults (≥ 18 years) hospitalized within past 10 years for complicated mild to severe TBI by GCS score of 3 to 12, or documented intracranial abnormalities on imaging, or post-traumatic amnesia of ≥ 7 days and meeting criteria for MDD on SCID and PHQ-9 scores ≥ 10, residing in the United States Exclusion criteria: no stable home or regular access to a telephone, history of schizophrenia, bipolar disorder, psychosis, suicidal intent, current alcohol or drug dependence, receiving or planning to receive psychotherapy for depression,	CBT-T x 8 sessions over 12 weeks (n=40)	CBT-IP x 8 sessions over 12 weeks (n=18) Usual care (n=42)	Primary: Change in depression severity on clinician-rated HAMD-17 and patient-reported SCL-20 over 16 weeks Secondary: -Major depressive disorder criteria based on SCID -PGI -Patient satisfaction -Therapeutic alliance -Quality of life -Functional

First Author, Year, Country	Study Design (Follow-up)	Study Population	Intervention (Sample size)	Comparator(s) (Sample size)	Reported Outcomes
		antidepressant plan to or initiate or recent dose adjustment, severe cognitive impairment			impairment by the Sheehan Disability Scale -Post-concussive symptoms
Anxiety					
Rollman, 2017[24] United States	Prospective, multicentre, assessor-blinded RCT (24 months)	Inclusion criteria: Adults (18 to 64 years old) meeting criteria for PD and/or GAD screened for anxiety disorders using the PRIME-MD anxiety module, reliable telephone access and a moderately severe level of anxiety as assessed by the SIGH-A (score ≥ 20) and PDSS (score ≥ 14) Exclusion criteria: active treatment with a mental health specialist, psychotic illness, alcohol dependence as determined by the AUDIT-C (score ≥ 3) and Rapid Alcohol Problems Screen	Telephone-delivered stepped collaborative care (n=124 high anxiety patients) Patients who were assessed as having moderate anxiety (n=79) were randomized to watchful waiting and upon deterioration to high anxiety; n=11 were randomized to the intervention)	Usual care (n=126 high anxiety patients) Of those randomized to the watchful waiting group, n=12 were randomized to usual care)	Primary: HRQL as measured by the SF-36 MCS within African-Americans at 12 months Secondary: -SIGH-A -PDSS -GADSS -PHQ-9
Depression and Anxiety					
van den Berg, 2015[25] Germany	Prospective, multicentre, open-label, 3-arm RCT (6 months)	Inclusion criteria: Adult patients preparing for discharge from psychiatric day hospital after 6 to 8 weeks treatment, diagnosis of depression, anxiety disorder, adjustment disorder, or somatoform disorder Exclusion criteria: interval patients (defined as patients scheduled to return to hospital after 3 to 6 months for continuation of therapy, distinct emotional instability with recurrent suicide crises and/or self-injuring beahvior	2 different telemedical interventions: Group 1: proactive, regular telephone contacts once per week x 1 month, then once per month x 5 months (n=42) Group 2: As per Group 1 plus weekly SMS text messages tailored to the individual patient (n=40)	Usual care (access health services as needed and desired) (n=41)	Primary: Severity of anxiety, depression, and somatization by the symptom scales of the German version of the BSI-18 No secondary outcomes stated although an exploratory sensitivity analysis restricted to 75% of participants with the highest symptom scores at baseline was conducted.
Post-Traumatic Stress Disorder					

First Author, Year, Country	Study Design (Follow-up)	Study Population	Intervention (Sample size)	Comparator(s) (Sample size)	Reported Outcomes
Franklin, 2017[26] United States	Prospective, multicentre, open-label, pilot RCT (6 weeks)	Inclusion criteria: Adult veterans meeting PTSD criteria or significant subclinical PTSD symptoms from military-related trauma as per the DSM-IV, impaired sleep in last month as per CAPS, rural (defined by living ≥ 30 miles from the PTSD clinic) Exclusion criteria: uncontrolled primary sleep disorders (e.g., obstructive sleep apnea), active psychosis, substance dependence, not enrolled in other active psychotherapies	CBT-T x 6 sessions in 8 weeks (n=11)	CBT-IP x 6 sessions in 8 weeks (n=7)	Primary: Insomnia by the self-reported PSQI at 6 weeks Secondary: -acceptability by an investigator-developed questionnaire -feasibility by number of referrals, rates of entry, eligibility, attrition from treatment

AUDIT-C = Alcohol Use Disorders Identification Test; BAI = Beck Anxiety Inventory; BATD = behavioral activation treatment for depression; BDI-II: Beck Depression Inventory II; BHS = Beck Hopelessness Scale; BSI-18 = Brief Symptom Inventory Short Version; CBT = cognitive behavioral therapy; CBT-IP = CBT in person; CBT-T = CBT by telephone; CIS-R = Clinical Interview Schedule-Revised Scale; CGI-I: Clinical Global Impressions - Improvement Questionnaire; CGI-S = Clinical Global Impressions - Severity of Illness Questionnaire; CSQ = Client Satisfaction Questionnaire; DSM -IV= Diagnostic and Statistical Manual (of Mental Disorders) IV; eHEALS = eHealth Literacy Scale; EQ-5D-5L = EuroQoL 5 Dimension 5 Level; GAD = generalized anxiety disorder; GADS-7 =GAD 7-item Scale; GADSS = Generalized Anxiety Disorder Severity Scale; GCS = Glasgow Coma Scale; GDS = Geriatric Depression Scale; HAMD-17 = 17-item Hamilton Depression Rating Scale; HeiQ = Health Education Impact Questionnaire; HRQL = health-related quality of life; IASMHS = Inventory of Attitudes Tow ard Seeking Mental Health Services; MINI = Mini-International Neuropsychiatric Interview; MST = military sexual trauma; Q-LES-Q: Chinese-translated version of the 'short-form' of the Quality of Life Enjoyment and Satisfaction Questionnaire; PCL-M = Post-Traumatic Stress Disorder Checklist - Military; PD = panic disorder; PDSS = Panic Disorder Severity Scale; PHQ-9 = Patient Health Questionnaire 9; PRIME-MD = Primary Care Evaluation of Mental Disorders Anxiety Module; PSQI = Pittsburgh Sleep Quality Index; RCT = randomized controlled trial; SCID = Structured Clinical Interview for the DSM-IV Axis I Disorders; SCL-20 = Symptom Checklist -20 Depression Scale; SF-36 MCS = Medical Outcomes Study Short-Form 36 Mental Component Summary; SIGN-A = Hamilton Anxiety Rating Scale; TAU = treatment as usual; TBI = traumatic brain injury; T-CSCT = telepsychiatry-based culturally sensitive collaborative treatment

Appendix 3: Critical Appraisal of Included Publications

Table 4: Strengths and Limitations of Systematic Reviews and Meta-Analyses using the AMSTAR 2 Checklist[12]

Strengths	Limitations
Depression	
Linde, 2015[44]	
-An 'a priori' design was described and a published protocol[27] predated the systematic review -A comprehensive literature search was performed of 5 databases (MEDLINE, Embase, Cochrane Central Register of Controlled trials [CENTRAL] and Psyc-INFO from 1980 onwoards -Literature searches were conducted in June 2011 with the last update in December 2013 -An explanation for the selection of study design was provided -A list of the included studies was provided -Characteristics of the included studies were provided -Study selection was done by 4 reviewers and data extraction was independently done by at least 2 reviewers using a pre-tested form -Quality of the included studies was assessed and the results were reported -Risk of bias was assessed using a suitable technique and overall risk of bias was considered low in 10 studies, unclear in 9 studies, and high in 11 studies -The methods used to statistically combine study results in the meta-analysis were appropriate -Statistical heterogeneity of study findings was investigated using various methods -Conflict of interest declaration was included -Source of funding for the systematic review was disclosed	-A list of excluded studies was not provided -Data on the source of funding for individual studies was not provided -The number of trials and sample sizes for some subgroups analyses were small therefore results are associated with uncertainty -The diagnosis of conditions in included studies (e.g., major depression, minor depression, dysthymia) was done using different methods in the included studies such as screening of primary care patients vs. physician referral -One third of included trials were considered to have a low risk of bias -Reporting of treatment discontinuation, dropouts, and loss to follow-up did not provide sufficient detail to determine the cause of attrition -None of the included studies reported AEs
Depression and Anxiety	
Coughtrey, 2018[43]	
-An 'a priori' design was described -A comprehensive literature search was performed of 3 databases (Psyc-INFO, PubMed and Web of Science) over a 25-year time period (January 1991 to May 2016) -The dates the literature searches were performed was not provided -An explanation for the selection of study design was provided -A list of the included studies was provided -Characteristics of the included studies were provided -Data extraction was done by the first author and the second author reviewed data extraction -Quality of the included studies was assessed and reported revealing that studies were generally strong in 3 domains: confounders, data collection methods, and withdrawals/dropouts -Only risk of selection bias was assessed in individual studies - Study heterogeneity was discussed but not quantified -Conflict of interest declaration was included	-The process by which study selection was done was not reported -A list of excluded studies was not provided -Data on the source of funding for individual studies was not provided -Results of the quality assessment revealed studies were weaker or more variable for the domains of: selection bias, design, and blinding. -Reporting of recruitment and dropout rates was variable and a number of studies had limited follow-up or did not follow-up post-intervention -Large heterogeneity in the included studies due to the broad inclusion criteria which precluded drawing firm conclusions regarding the effectiveness of the intervention

Strengths	Limitations
-Source of funding for the systematic review was disclosed; however, it was stated that the author(s) did not receive any financial support for the research, authorship, and/or publication of this review.	
Post-Traumatic Stress Disorder	
Turgoose, 2017[10]	
-An 'a priori' design was described -A comprehensive literature search was performed of 3 databases (CINAHL, PubMed, and Psyc-INFO), the Cochrane Library, and select peer-reviewed telemedicine journals (there were no restrictions in terms of publication date) -The dates that the literature searches were conducted were not provided -A list of the included studies was provided -Characteristics of the included studies were provided -Quality of the included studies was assessed for 28 studies with some form of experimental design as the tool used was not designed for studies with non-experimental designs; results indicated that 11 studies were strong, 8 were moderate, and 8 were considered weak -Risk of bias was assessed and studies that were considered weak had issues with selection bias, not controlling for potentially confounding variables and had large dropouts or did not report dropouts -Conflict of interest declaration was included -Source of funding for the systematic review was disclosed	-An explanation for the selection of study design was not provided -The processes for study selection and data extraction were not described -Study heterogeneity was not investigated -A list of excluded studies was not provided -Data on the source of funding for individual studies was not provided -Literature was dominated by studies from North America (all studies were from the United States with the exception of one Canadian study which has the potential to affect generalizability of the review's findings
Goetter, 2015[10]	
-An 'a priori' design was described -A comprehensive literature search was performed of 2 databases (PubMed and and Psyc-INFO (the date range for publications was not provided) -The literature searches were conducted between January and March 2014 -A list of the included studies was provided -Characteristics of the included studies were provided -Study selection was done by 2 reviewers -The methods used to statistically combine study results in the meta-analysis were appropriate -Statistical heterogeneity of study findings was investigated	-An explanation for the selection of study design was not provided -It was not clear if data extraction was done by 2 reviewers -A list of excluded studies was not provided -Quality of the individual studies and risk of bias was not assessed -Conflict of interest declarations were not provided -Source of funding for the systematic review was not disclosed -Data on the source of funding for individual studies was not provided -Included studies used a range of definition of dropout which could impact construct validity -All the studies were conducted within Veterans Affairs or Department of Defense clinics and in participants with combat-related PTSD which may affect generalizability to other clinic settings or PTSD populations

AE = adverse events; PTSD = Post-Traumatic Stress Disorder

Table 5: Strengths and Limitations of Included Randomized Controlled Trials using the Downs and Black Checklist[13]

Strengths	Limitations
Depression	
Heckman, 2017[17]	
-Objectives of the study were clearly described -The intervention and main outcomes were clearly described -Screening criteria for study eligibility were described -Study patients in the intervention and comparator groups were recruited from the same population, using the same inclusion criteria, over the same period -Study patients were randomized to treatment -Patient characteristics were clearly described and it was stated that there was no difference by treatment condition for any pre-intervention variable -Appropriate measures of random variability were reported -Main findings were clearly described -Appropriate statistical tests were used to assess the main study outcomes -Actual probability values were reported -Valid and reliable main outcome measures were used -Analyses were by ITT and a completer analysis -It was stated that no AEs were reported during the trial -The study setting was appropriate	-Patient flow was described but the number of patients lost to follow-up was not reported -Included patients were diagnosed with MDD as per DSM-IV criteria; however, all patients had AIDS or HIV infection which limits generalizability of the findings of this study to those without HIV/AIDS -Patients were not blinded to treatment; however teletherapists and the investigator conducting the randomization were blinded to pre-intervention data
Luxton, 2016[18]	
- Objectives of the study were clearly described -A protocol for the study was previously published[28] -The intervention and main outcomes were clearly described -Screening criteria for study eligibility were described -Study patients in the intervention and comparator groups were recruited from the same population, using the same inclusion criteria, over the same period -Study patients were randomized to treatment -Patient characteristics were clearly described and it was stated that there was no difference by treatment condition for any pre-intervention variable -Appropriate measures of random variability were reported -Main findings were clearly described -Appropriate statistical tests were used to assess the main study outcomes -The non-interiority margin was identified as a standardized difference of 0.50; the null hypothesis was to be rejected if the CI calculated around the treatment difference falls within the set threshold -Actual probability values were reported -Valid and reliable main outcome measures were used -Analyses were by ITT and a completer analysis -It was stated that no AEs were reported during the trial -Patients lost to follow-up were reported -The study setting was appropriate	-Patients and treatment providers were not blinded to the treatment allocation; however, clinician assessors who conducted outcome assessments were blinded to treatment condition -Sample size was insufficient to address superiority of the intervention -The study was not powered to make any inferences regarding safety
Salisbury, 2016[19]	
-Objectives of the study were clearly described	-The study recruitment was low so it is not known if patients with

Strengths	Limitations
-A protocol for the study was previously published[29] -The intervention and main outcomes were clearly described -Screening criteria for study eligibility were described -Study patients in the intervention and comparator groups were recruited from the same population, using the same inclusion criteria, over the same period -Study patients were randomized to treatment -Patient characteristics were clearly described and it was stated that the groups were mostly balanced -Appropriate measures of random variability were reported -Main findings were clearly described -Appropriate statistical tests were used to assess the main study outcomes -Actual probability values were reported -Valid and reliable main outcome measures were used -The primary analysis was based on ITT -All serious AEs were investigated -The source of funding for the study was described -Conflicts of interest were described by the authors -Patients lost to follow-up were reported -The study setting was appropriate	different characteristics would have received greater or lesser; benefit than was observed in the study -Once study eligibility was confirmed, the participation rate was high (84%) -Although groups were mostly balanced there were small baseline imbalances in work status, education, accommodation, depression severity, and use of antidepressants -Use of the intervention varied considerably; only 86 (28%) of patients in the intervention group received at least 9 of the potential maximum of 10 encounters -Patients were aware of their allocation but analysis of outcomes was masked.
Yeung, 2016[20]	
-Objectives of the study were clearly described -The intervention and main outcomes were clearly described -Screening criteria for study eligibility were described -Study patients in the intervention and comparator groups were recruited from the same population, using the same inclusion criteria, over the same period -Study patients were randomized to treatment -Patient characteristics were clearly described and it was stated that there were no baseline differences in the distributions of a number of key baseline variables -Appropriate measures of random variability were reported -Main findings were clearly described -Appropriate statistical tests were used to assess the main study outcomes -Actual probability values were reported -Valid and reliable main outcome measures were used -The primary analysis was based on ITT -AEs were monitored over the study period -The source of funding for the study was described -Conflicts of interest were declared by the authors -The study setting was appropriate	-Patient disposition after randomization was not provided and no information on patients lost to follow-up was reported -Patients were aware of their allocation however outcomes assessors/interviewers over the telephone were blinded to treatment -The study population was Chinese-American immigrants receiving primary care at a community health centre so it is unknown if the study results can be generalized to other populations
Egede, 2015[21,22]	
- Objectives of the study were clearly described -The intervention and main outcomes were clearly described -Screening criteria for study eligibility were described -Study patients in the intervention and comparator groups were recruited from the same population, using the same inclusion criteria, over the same period -Study patients were randomized to treatment -Patient characteristics were clearly described and it was stated that there was no difference by treatment condition for any pre-	-Patients were aware of their allocation however outcomes assessors/interviewers were blinded to treatment -The treatment response was less than anticipated when the study was designed which the authors speculate may be due to a sicker and more severely depressed patient population with psychiatric comorbidity than estimated -The technology used in the study was considered somewhat obsolete at the time of study completion -The study population was almost all male (98%) which limits

Strengths	Limitations
intervention variable -Appropriate measures of random variability were reported -Main findings were clearly described -Appropriate statistical tests were used to assess the main study outcomes -The non-interiority margin was based upon an absolute difference in the proportions for the primary outcome of 15%; the null hypothesis was to be rejected if the 90% CIs of the difference exceeded the set threshold (0.15) -Actual probability values were reported -Valid and reliable main outcome measures were used -Analyses were by ITT -It was stated that no AEs for any participant were noted in the study -Patients lost to follow-up were reported -The funding source for the study was reported -Conflicts of interest were declared by the authors -The study setting was appropriate	generalizability to female patients
Fann, 2015[23]	
- Objectives of the study were clearly described -The intervention and main outcomes were clearly described -Screening criteria for study eligibility were described -Study patients in the intervention and comparator groups were recruited from the same population, using the same inclusion criteria, over the same period -Study patients were randomized to treatment -Patient characteristics were clearly described and potential confounders and distributions were discussed -Appropriate measures of random variability were reported -Main findings were clearly described -Appropriate statistical tests were used to assess the main study outcomes -Actual probability values were reported -Valid and reliable main outcome measures were used -Analyses were by ITT -It was stated that there were no study-related AEs -Patient disposition was reported but it was unclear if patients reported as 'unable to contact' were equal to lost to follow-up -The funding source for the study was reported -Conflicts of interest were declared by the authors -The study setting was appropriate	- Included patients had a diagnosis of MDD as per PHQ-9 and SCID scores; however, all patients had history of traumatic brain injury which limits generalizability of the findings of this study to those without traumatic brain injury -Included study patients had high levels of medical, cognitive, and psychiatric comorbidity, which coupled with the small sample size, may have affected the ability to demonstrate statistical significance for some outcomes
Anxiety	
Rollman, 2017[24]	
-Objectives of the study were clearly described -The intervention and main outcomes were clearly described -Screening criteria for study eligibility were described -Study patients in the intervention and comparator groups were recruited from the same population, using the same inclusion criteria, over the same period -Study patients were randomized to treatment -Patient characteristics were clearly described and potential confounders and distributions were discussed	-Although the screening procedure was described, it was not a systematic screen of patients for a mental health disorder, but rather was based on an electronic medical record-generated alert to identify patients for treatment which limited recruitment to settings with systems capable of generating the alert and to proper entry of the diagnostic or medication codes into the alert -A large proportion of patients did not complete some study activities (e.g., workbook) which may have limited effectiveness -No safety outcomes were reported

Strengths	Limitations
-Appropriate measures of random variability were reported -Main findings were clearly described -Appropriate statistical tests were used to assess the main study outcomes -Actual probability values were reported -Valid and reliable main outcome measures were used -Analyses were by ITT -Patients disposition was reported but it was unclear if patients reported as 'unable to contact' were the same as lost to follow-up -The funding source for the study was reported -Conflicts of interest were declared by the authors -The study setting was appropriate	
Depression and Anxiety	
van den Berg, 2015[25]	
-Objectives of the study were clearly described -A protocol for the study was previously published[30] -Screening criteria for study eligibility were described -Study patients in the intervention and comparator groups were recruited from the same population, using the same inclusion criteria, over the same period -Study patients were randomized to treatment -Patient characteristics were clearly described -Appropriate measures of random variability were reported -Main findings were clearly described -Appropriate statistical tests were used to assess the main study outcomes -Actual probability values were reported -Valid and reliable main outcome measures were used -Analyses were by ITT -Patients lost to follow-up reported -The funding source for the study was reported -Conflicts of interest were declared by the authors -The study setting was appropriate	-The broad inclusion criteria led to the study population having high variability in the severity of patients' mental health disorders and comorbidities which may have reduced the potential effect of the intervention, although this strengthens external validity of the findings -Patients were recruited after treatment in a psychiatric day hospital so their clinical symptoms were well controlled at study entry -Potential confounders and distributions were not discussed -Patients and assessors/interviewers were not blinded to treatment -No safety outcomes were assessed
Post-Traumatic Stress Disorder	
Franklin, 2017[26]	
-Objectives of the study were clearly described -A protocol for the study was previously published[30] -Screening criteria for study eligibility were described -Study patients in the intervention and comparator groups were recruited from the same population, using the same inclusion criteria, over the same period -Study patients were randomized to treatment -Patient characteristics were clearly described -Appropriate measures of random variability were reported -Main findings were clearly described -Appropriate statistical tests were used to assess the main study outcomes -Actual probability values were reported -Valid and reliable main outcome measures were used -Analyses were conducted only in patients with complete data -Patients lost to follow-up reported	-As this was a pilot study, the study sample is very small (N=18) and inadequately powered to draw any conclusions regarding equivalence or superiority of the intervention versus the comparator -The patient population was entirely male which precludes generalizability of results to female military personnel -Analysis by ITT was not conducted -Functional outcomes were not assessed (only sleep outcomes) -No safety outcomes were assessed

Strengths	Limitations
-The funding source for the study was reported -Conflicts of interest were declared by the authors -The study setting was appropriate	

AE = adverse event; AIDS = acquired immunodeficiency syndrome; CI = confidence interval; DSM -IV= Diagnostic and Statistical Manual (of Mental Disorders) IV; HIV = human immunodeficiency virus; ITT = intention-to-treat; MDD = major depressive disorder; PTSD = post-traumatic stress disorder

Appendix 4: Main Study Findings and Authors' Conclusions

Table 6: Summary of Findings of Included Systematic Reviews and Meta-Analyses

Main Study Findings	Authors' Conclusions
Depression	
Linde, 2015[14]	
Findings for outcomes of interest (with number of studies and sample sizes) in patients with depression classified by type of treatment. *Post-treatment depression scores: SMD for mean score in treatment (T) minus mean score in control (C) divided by pooled SD [SMD (95% CI); P-value]:* *(Negative SMDs indicate a better outcome in the treatment group)* Face-to-face CBT (7 studies): T (n=251); C (n=267): -0.30 (-0.48; -0.13); *P*=0.0007 Face-to-face PS (4 studies): T (n=304); C (n=297): -0.14 (-0.40; 0.120); *P*=0.30 Face-to-face IPT (2 studies): T (n=151); C (n=154): -0.24 (-0.47; -0.02); *P*=0.04 Other face-to-face therapies (5 studies): T (n=534); C (n=415): -0.28 (-0.44; -0.12); *P*=0.0005 Remote therapist-led CBT (3 studies): T (n=335); C (n=342): -0.43 (-0.62; -0.24); *P*=0.00001 Remote therapist-led PS (2 studies): T (n=16); C (n=22): -0.56 (-1.57; 0.45); *P*=0.28 Differences between effects of all psychological interventions: *P*=0.74 *Pooled estimates for response: OR (95% CI); NNT (95% CI)* *(OR greater than 1 indicates more events in the treatment group versus the control group)* Face-to-face CBT (7 studies): 1.58 (1.11; 2.26); 10 (5; 47) Face-to-face PS (3 studies): 1.56 (0.85; 2.86); NA Face-to-face IPT (2 studies): 1.28 (0.80; 2.05); NA Other face-to-face therapies (6 studies): 1.54 (1.17; 2.03); 11 (6; 31) Remote therapist-led CBT (3 studies): 2.04 (1.44; 2.90); 6 (4; 13) Remote therapist-led PS (2 studies): No data *Pooled estimates for remission: OR (95% CI); NNT(95% CI)* Face-to-face CBT (7 studies): 1.49 (0.90; 2.46); NA Face-to-face PS (4 studies): 1.29 (0.83; 2.02); NA Face-to-face IPT (2 studies): 1.37 (0.81; 2.34); NA Other face-to-face therapies (5 studies): 1.68 (1.17; 2.41); 10 (5; 35) Remote therapist-led CBT (3 studies): 1.51 (0.98; 2.32)); NA Remote therapist-led PS (2 studies): 1.22 (0.23; 6.57); NA *Pooled estimates for study discontinuation: OR (95% CI)* Face-to-face CBT (6 studies): 0.98 (0.52; 1.86) Face-to-face PS (4 studies): 0.60 (0.23; 1.57) Face-to-face IPT (1 study): 0.98 (0.40; 2.38) Other face-to-face therapies (5 studies): 1.04 (0.65; 1.67) Remote therapist-led CBT (2 studies): 1.05 (0.29; 3.75) Remote therapist-led PS (2 studies): 1.32 (0.33; 5.26) *Subgroup analysis by delivery mode [SMD (95% CI); No. of comparisons]:* Face-to-face: -0.23 (-0.33; -0.14); 18 Remote therapist-led: -0.43 (-0.62; -0.25); 5 Test for subgroup differences: *P*=0.14	The authors concluded that *"For CBT approaches, substantial evidence suggests that interventions that are less resource intensive might have effects similar to more intense treatments."*[14] p56 The authors also stated *"Our analyses suggest that the differences between different types of psychological treatments are minor, and remote therapist-led, guided self-help, and minimal-contact approaches can yield effects similar to personalized face-to-face therapies."*[14] p65
Depression and Anxiety	

Main Study Findings	Authors' Conclusions
Coughtrey, 2018[15]	
Depression studies *RCTs:* 5 of 6 included RCTs reported statistically significant reductions following T-CBT (3 studies) or IPT (2 studies) vs. no T, TAU, enhanced usual care, or alternative active T; 1 RCT reported T-CBT did not significantly reduce depressive symptoms vs. TAU. Effect sizes could be analyzed for 5 of 6 RCTs: Cohen's *d* ranged from 0.25-1.98 (median=0.58) *Quasi-experimental and uncontrolled studies:* One quasi-experimental study using a benchmark comparison with published data found statistically significant reductions in depression after T-CBT. 3 uncontrolled studies reported statistically significant reductions in depression after T-CBT. Effect sizes could be analyzed for 2 of 3 studies: Cohen's *d* ranged from 1.13-1.90 (median=1.25). *Anxiety studies* *RCTs*: All 3 included RCTs reported statistically significant reductions in anxiety symptoms following the telephone-delivered intervention. Effect sizes could be analyzed for 2 of 3 RCTs and Cohen's *d* ranged from 0.34-1.07 (median=0.69). *Quasi-experimental study:* One controlled, non-randomized study compared telephone delivered CBT and ERP with a delayed wait-list control group and found statistically significant reductions in symptoms compared with the control (Cohen's *d* 1.07).	The authors concluded that *"The findings indicate that telephone-delivered interventions show promise in reducing signs of depression and anxiety. Further research is required to establish the types of interventions that are most effective and the characteristics of clients who find them beneficial."*[15] p65
Post-Traumatic Stress Disorder	
Turgoose, 2017[10]	
Treatment outcomes 15 of 18 studies that evaluated clinical effectiveness of tele-therapy vs. a control (typically the same intervention but in-person) reported tele-therapy was associated with significant reductions in PTSD symptoms, regardless of the type of intervention, except for 1 study that only measured anger in veterans with PTSD. 12 of the 18 studies reported comparisons between tele-therapy and in-person interventions using non-inferiority analyses. Of these, 9 studies concluded tele-therapy was as effective as in-person therapy, 2 studies suggested in-person interventions had significantly greater reductions in PTSD symptoms, and 1 study found tele-therapy to be more effective than usual care. *Dropout rates* 13 studies reported findings related to attrition, dropout, and attendance rates. No significant differences in attrition between tele-therapy and in-person treatments were found in any of the studies, although 1 study found those receiving tele-therapy attended significantly more sessions. Individual studies suggested differences in treatment uptake but in the majority of studies there were no differences between tele-therapy and in-person treatments on different process variables. *Acceptability* No studies found significant differences in patient satisfaction and acceptability between tele-therapy and in-person treatment groups, with most reporting high levels of satisfaction for both. *Therapeutic alliance* 10 studies investigated the quality of therapeutic alliance and compared tele-therapy and in-person interventions with mixed results. 5 of 6 studies of one-to-one treatment reported equivalence in therapeutic alliance between tele-therapy and in-person treatment whereas 1 study reported in-person treatment to be more comfortable. 3 studies reported on therapeutic alliance in groups of which 1 study found no differences between modalities, 2 studies found lower group therapeutic alliance or lower alliance with the group leader. *Managing risk and safety* 4 studies reported on issues relating to suicidality or general patient safety in the context of high	The authors concluded that *"Early findings suggest that tele-therapy is as effective as in-person treatments, and comparable in terms of acceptability and therapy process issues, all of which suggest it is a viable alternative to in-person treatments. However, there are some instances of differences between tele-therapy and in-person treatments, such as therapeutic alliance. Therefore, tele-therapy should continue to be evaluated and scrutinized in order to ensure it is used in a way that maximizes its effectiveness."*[10] p9

Main Study Findings	Authors' Conclusions
emotional levels in trauma-focused therapies. Of these 3 studies reported no significant difficulties in using tele-therapy to manage the situation whereas 1 study, while not reporting statistically significant differences, reported a trend for participant concern about managing emotions during exposure tasks via tele-therapy (i.e., without the physical presence of the therapist). *Feasibility* 10 studies described the use of technology in tele-therapy interventions and commonly reported technical difficulties were low image resolution, difficulty connecting, and slight audio delays but overall did not adversely affect the therapy.	
Goetter, 2015[16b]	
Findings from 20 studies (n=1,191) of veterans (81.7%) or active duty service members (18.3%) with combat-related PTSD with mean age (based on available data) ranging from 25.8 to 39.0 years and proportion of females ranging from 0.0% to 20.0%: *Dropout rates (95% CI):* Overall pooled drop-out rate: 36.0% (26.90; 45.00) - by study type: OR:0.86 (0.73; 1.02); *P*=0.072; reported average dropout rates were 42.0% (routine clinical care setting) and 28.0% (clinical trial) -by treatment format: OR: 1.28 (1.06; 1.53); *P*=0.012; reported average drop rates were 31.1% (25.10; 37.20) for individual therapy and 54.4% (30.20; 78.60) for group therapy component -by exclusion of substance dependence: OR:1.04 (0.86; 1.25) -by treatment modality (telemedicine vs. in-person treatment): OR: 0.96 (0.72; 1.28) -by treatment type (exposure therapy vs. non-exposure therapy): OR:1.00 (0.86; 1.18) Identified correlates of dropout were age (more likely to be younger than completers), being unmarried, unemployed, having service-connected disability, and greater PTSD severity	The authors stated that *"In conclusion, dropout rates varied considerably, but approximately one third of OEF/OIF/OND veterans with combat-related PTSD drop out of PTSD treatment. Novel interventions are needed to mitigate the problem of dropout."*[16] p406

C = control; CBT = cognitive behavioral therapy; CI = confidence interval; ERP = exposure and response prevention; IPT = interpersonal psychotherapy; NNT = number needed to treat; OR = odds ratio; PS = problem solving; PTSD = post-traumatic stress disorder; RCT = randomized controlled trial; SD = standard deviation; SMD = standardized mean difference; T = treatment; TAU = treatment as usual; T-CBT = telephone-delivered CBT

Table 7: Summary of Findings of Included Randomized Controlled Trials

Main Study Findings	Authors' Conclusions
Depression	
Heckman, 2017[17]	
Baseline characteristics: -Total study population (N=132); mean age: 51.9 years; 63% male; 75% Caucasian; 57% diagnosed with AIDS; 54% self-identified as gay/bisexual; 81% diagnosed with MDD and remainder with partial remission of MDD or dysthmic disorder; 62% prescribed antidepressant therapy. Mean (SD) pre-intervention BDI score was 26.6 (9.4), IIP score was 100.9 (33.1), and PSRS score was 40.8 (9.3). There were no significant differences between groups on any pre-intervention variables. *Dropout rate:* 14% did not complete post-intervention assessments (n=13 T-IPT vs. n=6 standard care) *Changes from pre-intervention through post-intervention by study arm, mean (SD):* BDI-II (Pre/Post/Difference): MMRM imputation: T-IPT (n=70): 26.3 (1.0)/20.9 (1.3)/-5.4 Standard care (n=62): 27.0 (1.3)/25.4 (1.5)/-1.6 Completer only: T-IPT (n=56): 26.0 (1.3)/20.2 (1.4)/-5.8 Standard care (n=62): 27.0 (1.2)/25.6 (1.4)/-1.4 A significant "time x study arm" interaction was found for the MMRM (*P*=0.012) and Completer only (*P*=0.005) analyses IIP (Pre/Post/Difference): MMRM imputation: T-IPT (n=70): 104.4 (4.0)/93.0 (4.8)/-11.4 Standard care (n=62): 97.1 (4.1)/98.5 (4.5)/+1.4 Completer only: T-IPT (n=56): 104.8 (4.9)/90.1 (5.3)/-14.7 Standard care (n=62): 97.5 (4.9)/100.7/+3.2 A significant "time x study arm" interaction was found for the MMRM (*P*=0.006) and Completer only (*P*=0.002) analyses PSRS (Pre/Post/Difference): MMRM imputation: T-IPT (n=70): 41.0 (1.1)/40.1 (1.4)/-0.90 Standard care (n=62): 40.7 (1.2/40.1 (1.3)/-0.60 Completer only: T-IPT (n=56): 40.2 (1.2)/39.4 (1.4)/-0.80 Standard care (n=62): 40.7 (1.2)/40.1 (1.4)/-0.60 PSRS scores were unchanged for both the MMRM (*P*=0.860) and Completer only (*P*=0.900) analyses Note: ITT analyses used data from all patients with pre-intervention BID values ≥ 14. Completer only analyses used data from T-IPT patients who attended all 9 tele-therapy sessions as well as all standard care controls. Results of sensitivity analyses conducted using data from all patients enrolled (N=167) regardless of pre-intervention BDI using ITT and Completer only analyses did not differ from the above analyses. No "gender x condition" interaction was found. Responders were defined as a reduction of 50% or more in depressive symptoms based on the BDI. Results for non-responders/partial responders/responders were: T-IPT: 50%/28%/22% vs. Standard care: 78%/18%/4%, respectively, for the MMRM multiple imputation analysis and T-IPT: 48%/28%/23% vs. Standard care: 82%/15%/3%, respectively, for the Completers only analysis. A significant "study arm x treatment response" was found for both the MMRM (*P*<0.004) and Completer only (*P*<0.001) analyses.	The authors concluded that *"Brief tele-IPT acutely decreased depressive symptoms and interpersonal problems in depressed rural people living with HIV"*[17] p285

<table>
<tr><th>Main Study Findings</th><th>Authors' Conclusions</th></tr>
<tr><td>No adverse events were reported during the trial.</td><td></td></tr>
<tr><td colspan="2">Luxton, 2016[18]</td></tr>
<tr><td>Baseline characteristics
-Total study population (N=121); 61% of patients ≤ 34 years old; 82% male; 70% Caucasian
-In-person group: 54 (91.53%) and In-home group: 56 (90.32%) met SCID-I/P criteria for MDD

Dropout rate: 32.23% (n=40 in-home and n=42 in-office completed all 8 sessions)

Primary outcome: BHS and BDI-II scores post-treatment, mean reduction (95% CI); ITT analysis
-BHS: In-person group: -6.21 points (-7.38; -5.50) vs. In-home: -3.91 points (-5.25; -2.57)
-BDI-II: In-person group: -17.63 (-20.21; -15.06) vs. In-home: -13.40 (-16.36; -10.44)
-After standardization of the difference between groups at post-assessment, the upper bound of the 90% CI fell below the pre-specified non-inferiority margin of 0.50 supporting that in-home therapy was non-inferior to In-person therapy
-The per-protocol analysis provided similar results

Selected secondary outcomes:
-At post-treatment 8 (17.78%) In-home and 6 (14.29%) In-person met criteria for MDD and difference was not statistically significant
-Participants in both treatment groups reported reductions in anxiety and post-traumatic stress symptoms and in mental health treatment stigma as measured by the IASMHS and average scores on the CSQ suggested a high level of treatment satisfaction for both treatments. There were no statistically significant differences between groups for these outcomes.

Safety outcomes:
-AEs were reported in 7 patients in the in-home group and 4 patients in the in-person group; none of the AEs (e.g., severe exacerbation of asthma symptoms) was deemed related to study procedures</td><td>The authors concluded that "BATD can be feasibly delivered to the homes of active duty members and veterans via VC. Small-group differences suggest a slight benefit of in-person care over in-home telehealth on some clinical outcomes."[18] p 923</td></tr>
<tr><td colspan="2">Salisbury, 2016[19]</td></tr>
<tr><td>Baseline characteristics
-Total study population (N=609):Mean age 49 to 50 years; 68% female; 97% Caucasian
-205 (34%) were severely depressed
-527 (92%) had been previously treated for depression
-509 (88%) were currently taking antidepressant medication

Dropout/exclusion rate at 4 months (primary outcome): 14% (n=52 intervention + usual care and n=32 usual care)

Primary outcome: Repeated measures analysis of treatment response[a] as a binary outcome
<table>
<tr><th></th><th>Usual care</th><th>Intervention + Usual care</th><th>Adjusted OR (95% CI)</th><th>P-value</th></tr>
<tr><td>4 months</td><td>50/270 (19%)</td><td>68/255 (27%)</td><td>1.7 (1.1; 2.5)</td><td>0.019</td></tr>
<tr><td>8 months</td><td>61/263 (23%)</td><td>75/252 (30%)</td><td>1.4 (1.0; 2.2)</td><td>NC</td></tr>
<tr><td>12 months</td><td>86/261 (33%)</td><td>95/255 (37%)</td><td>1.2 (0.9; 1.8)</td><td>NC</td></tr>
<tr><td>Average of 4, 8, and 12 months[a]</td><td>-</td><td></td><td>1.6 (1.0; 2.6)</td><td>0.035</td></tr>
</table>
Note: all analyses are adjusted by site and baseline PH-9 score; GP practice is included as a random effect
CI = confidence interval; NC=not calculated at individual timepoints; analysis plan specified repeated measures analysis; OR = odds ratio
[a] Response defined as PHQ-9 score ≤ 10 and a reduction of ≥ 5 points at 4 months after randomization
[b] Based on a repeated measures analysis that was additionally adjusted by follow-up timepoint as a categorical variable
Table used in full: Salisbury C, O'Cathain A, Edwards L, Thomas C, Gaunt D, Hollinghurst S, et al. Effectiveness of an integrated telehealth service for patients with depression: a pragmatic randomised controlled trial of a complex intervention. Lancet Psychiatry. 2016 Jun;3(6):515-25. Available from: http://www.thelancet.com/journals/lanpsy/article/PIIS2215-0366(16)00083-3/fulltext
Creative Commons Link: https://s100.copyright.com/AppDispatchServlet?publisherName=ELS&contentID=S2215036616000833&orderBeanReset=true&orderSource=Pho</td><td>The authors stated that the added value of this study was that "Our trial builds on previous evidence by testing a new model of care that combines the use of various telehealth tools supported by non-clinically trained staff. Our study provides robust support for the potential role of an integrated telehealth</td></tr>
</table>

Main Study Findings	Authors' Conclusions

enix

Selected secondary outcomes and measures of assessment

	Usual care		Intervention + usual care		Adj diff in means (95% CI)	*P*-value
	n	Unadj mean (SD)	n	Unadj mean (SD)		
PHQ-9 score:						
Baseline	302	16.7 (4.7)	307	17.1 (4.5)	-	-
12 months	261	11.9 (6.4)	255	11.6 (6.2)	-0.5 (-1.5; 0.5)	NC
Av of 4, 8, and 12 months[a]	-	-	-	-	-0.8 (-1.6; 0.0)	0.045
GAD-7	237	9.2 (5.8)	223	8.7 (5.5)	-1.1 (-2.0; -0.2)	0.019
EQ-5D-5L	227	0.564 (0.30)	219	0.569 (0.30)	-0.003 (-0.04; 0.04)	0.896
Treatment satisfaction	184	3.3 (0.9)	193	3.7 (0.9)	-0.5 (0.3; 0.6)	<0.001
Difficulty accessing care	232	4.2 (1.9)	216	4.5 (1.9)	0.3 (0.0; 0.6)	0.038
Satisfaction with support	177	2.1 (0.9)	185	2.6 (0.8)	0.6 (0.4; 0.7)	<0.001

Note: all analyses are adjusted by site, baseline outcome (if measured) and baseline PH-9 score; GP practice is included as a random effect
Adj = adjusted; CI = confidence intervals; GAD - 7 = Generalized Anxiety Disorder 7 item scale; EQ-5D-5L = EuroQoL 5 Dimention, 5 Level; PHQ-9 = Patient Health Questionnaire-9; NC = not calculated at individual timepoints; analysis plan specified repeated measures; SD = standard deviation; Unadj = unadjusted

[a] Based on a repeated measures analysis that was additionally adjusted by follow-up timepoint as a categorical variable

Table Adapted from: Salisbury C, O'Cathain A, Edwards L, Thomas C, Gaunt D, Hollinghurst S, et al. Effectiveness of an integrated telehealth service for patients with depression: a pragmatic randomised controlled trial of a complex intervention. Lancet Psychiatry. 2016 Jun;3(6):515-25. Available from: http://www.thelancet.com/journals/lanpsy/article/PIIS2215-0366(16)00083-3/fulltext

Creative Commons Link:
https://s100.copyright.com/AppDispatchServlet?publisherName=ELS&contentID=S2215036616000833&orderBeanReset=true&orderSource=Phoenix

Safety outcomes
-70 AEs reported overall (36 in usual care and 34 in intervention + usual care groups)
-1 non-serious AE was related to the intervention (i.e., anxiety from talking about depression)
-2 deaths (1 due to throat cancer and 1 due to COPD) both in usual care group

approach to support patients with depression in primary care."[19] p516

Yeung, 2016[20]

Baseline characteristics:
-Total study population (N=190); mean (SD) age: 50 (14.5) years; 63% female
-Mean (SD) baseline scores were: HAMD17: 19.6 (4.2); CGI-S: 4.2 (0.8); Q-LES-Q total: 40.9 (6.6)

Dropout rate: not reported

Primary outcome: Multivariate analysis of the effects of T-CSCT (ITT)
Response, n/N (%):
Usual care: 16/94 (17.0) vs. Intervention: 41/92 (45.0); P < 0.001[a]; OR: 3.9 (95% CI: 1.9; 7.8); P <0.001[b]

Remission, n/N (%):
Usual care: 10/96 (10.4) vs. Intervention: 30/93 (32.3); P < 0.001[a]; OR: 4.4 (95% CI: 1.9; 9.9); P <0.001[b]

[a] $\chi 2$ test for TAU vs. T-CSCT
[b] Multivariate logistic regression for TAU vs. T-CSCT

Secondary outcomes:
Tests of "time x treatment" interaction for repeated measures found the T-CSCT group had significantly greater improvement over time than TAU for HAMD17 (P=0.002), CGI-S (P=0.003), and CGI-I (P=0.02) but there was no significant difference between groups for Q-LES-Q scores (P=0.82).

Safety outcomes: none reported.

The authors concluded that *"T-CSCT improved treatment outcomes of depressed Chinese-Americans in primary care. With the use of teleconferencing technologies, there is a potential for large-scale dissemination of CSCT to address disparities in the treatment of depression."*[20] p1001

Main Study Findings	Authors' Conclusions
Egede, 2015,[21] 2016[22]	
Baseline characteristics: -Total study population (N=241); mean age 63.9 years (5.1); 98% male; 60% Caucasian -Mental status questionnaire score: 67% normal mental function; 27% 1 error; 6% ≥ 2 errors -Baseline scores (SD): GDS: 20.8 (4.8); BDI: 26.8 (10.0) -Lifetime prevalence of GAD: 42%; PTSD: 63%; symptomatic diagnosis of PTSD in past month: 79% -Rural residents: 71% *Dropout rate:* 15.35% (83% of the telemedicine group and 86% of the same-room group returned for the final assessment) *Primary outcome: Treatment response*[a] *at 12 months* <u>ITT analysis:</u> BDI, n (% [90% CI]): Telemedicine: 27 (22.54% [15.40; 29.69]) vs Same-room: 26 (21.49% [14.72; 28.25]); Difference: 1.05% (90% CI: -8.30; 10.41) GDS, n (% [90% CI]): Telemedicine: 25 (20.96% [14.45; 27.47] vs. Same-room: 23 (19.30% [13.29; 25.31]); Difference: 1.66% (90% CI: -7.20; 10.52) SCID, n (% [90% CI]): Telemedicine: 53 (44.17% [35.78; 52.55]) vs. Same-room: 58 (47.85% [39.63; 56.07]); Difference: -3.68% (90% CI: -15.53; 8.16) Per-protocol analysis: BDI, n (% [90% CI]): Telemedicine: 19 (24.05% [16.14; 31.96]) vs Same-room: 19 (23.17% [15.51; 30.83]); Difference: -0.88% (90% CI: -10.13; 11.89) GDS, n (% [90% CI]): Telemedicine: 22 (22.45% [15.52; 29.38] vs. Same-room: 21 (20.39% [13.86; 26.92]); Difference: 2.06% (90% CI: -7.46; 11.58) SCID, n (% [90% CI]): Telemedicine: 39 (43.33% [34.74; 51.93]) vs. Same-room: 46 (48.42% [39.99; 56.85]); Difference: -5.09% (90% CI: -17.13; 6.95) [a] Defined as a 50% reduction in symptoms from baseline at 12 months For both BDI and GDS, the lower limits of the 90% CI do not exceed the pre-specified non-inferiority margin of 0.15 supporting that telemedicine was non-inferior to same-room treatment. The difference in treatment response based on SCID was not significant (*P*=0.487) No statistically significant differences were found between treatment groups in the mean scores over time for each of the 8 quality of life domains of the SF-36 in either the ITT or Per-protocol populations. Similarly, there were no statistically significant differences in patient satisfaction based on treatment credibility (as assessed by the Treatment Credibility Questionnaire) or the Charleston Psychiatric Outpatient Satisfaction Scale. Means scores for service delivery perception also did not show statistical differences for most of the variables with the exception of the same-room group having superior scores at study end in the categories of 'likelihood of return' and 'overall satisfaction' which became non-significant after a Bonferroni correction for multiple outcomes in both the ITT and Per-protocol populations. *Safety outcomes:* No AEs were reported for any study participant	In the first publication of this study, the authors concluded that *"Telemedicine-delivered psychotherapy for older adults with major depression is not inferior to sameroom treatment. This finding shows that evidence-based psychotherapy can be delivered, without modification, via home-based telemedicine, and that this method can be used to overcome barriers to care associated with distance from and difficulty with attendance at in-person sessions in older adults."*[21] p693 In the second publication of this study, the authors concluded that *"In conclusion this study found telemedicine-delivered behavioral activation treatment for elderly veterans with depression to be non-*

Main Study Findings	Authors' Conclusions
	inferior to face-to-face treatment in terms of quality of life, satisfaction, treatment credibility, and service delivery perception."[22] p1710
Fann, 2015[23]	
Baseline characteristics: -Total study population (N=100); mean (SD) age: 45.8 (13.3); 63% male; 90% non-Hispanic white -Randomization choice: CBT-IP vs. Usual care (4%); CBT vs. Usual care (47%), and CBT-IP vs. CBT-T vs. Usual care (49%); mean (SD) years since TBI: 3.33 (2.72); 18% had history of PTSD *Dropout rate (16 weeks):* CBT-T (7%), CBT-IP (17%), and UC (14%) *Primary outcome: Depression outcomes (ITT)* Results are: 'All CBT' (n=58)/'CBT-T' (n=40)/'CBT-IP' (n=18)/'UC' (n=42) HAMD17 score[a] Baseline: 17.5 (3.9)/17.5 (3.3)/17.5 (4.3)/17.6 (4.3) Week 16: 11.6 (6.1)/11.5 (6.2)/12.2 (6.8)/12.2 (6.8) Week 16 response[b]: 17 (33%)/12 (32%)/11 (31%)/11 (31%) SCL-20[a] Baseline: 1.99 (0.55)/2.07 (0.58)/1.81 (0.45)/1.89 (0.50) Week 16 1.18 (0.72)/1.39 (0.62)/1.18 (0.68)/1.30 (0.68) Week 16 response[b]: 23 (44%)/17 (46%)/6 (40%)/10 (28%) MDD negative on SCID[b] Week 16: 38 (73%)/26 (70%)/12 (80%)/20 (57%) Week 24: 35 (67%)/24 (69%)/11 (65%)/23 (68%) PGI - much or very much improved, n (%) Week 16: 32 (62%)/25 (68%)/7 (47%)/14 (39%) Week 24: 29 (58%)/23 (68%)/6 (38%)/14 (40%) Satisfaction - moderately or very satisfied, n (%) Week 16: 41 (84%)/30 (83%)/12 (86%)/9 (26%) [a] Mixed model regression using an autoregressive correlation matrix, reporting the significance of the time x treatment term [b] Response defined as ≥ 50% reduction from baseline; significance by exact logistic regression (using covariates from linear model) *Treatment comparisons of interest: CBT-T (n=40) vs. UC (n=39):* -Statistically significant differences were demonstrated for SCL-20 at Week 8 (*P*=0.002) and Week 16 (*P*=0.043), for PGI - much or very much improved at Week 16 (*P*=0.012) and Week 24 (*P*=0.026), and for Satisfaction with depression care - moderately or very satisfied at Week 16 (*P*<0.001) -There were no significant group differences on HRQL, functional impairment, overall post-concussive symptom score, environmental reward, automatic negative thoughts, or dysfunctional attitudes.	The authors concluded that *"In-person and telephone-administered CBT are acceptable and feasible in persons with TBI. Although further research is warranted, telephone CBT holds particular promise for enhancing access and adherence to effective depression treatment."* [23] p 45

Main Study Findings	Authors' Conclusions
Safety outcomes: No study-related AEs were reported	
Anxiety	
Rollman, 2017[24]	
Baseline characteristics: -Total study population (N=329); n=250 highly anxious patients and n=79 moderately anxious patients triaged to a watchful waiting cohort of which n=23 were later randomized due to deterioration; mean (SD) age: 44.6 (11.2)/42.1 (11.1); male: 65%/30%; Caucasian: 78 (194)/85 (67) for highly/moderately anxious patients -116 (46%) of highly anxious and 38 (48%) of moderately anxious patients had a diagnosis of GAD, 121 (48%) and 21 (27%) had a diagnosis of GAD and panic disorder, and 225 (90%) and 54 (68%) had a diagnosis of MDD, respectively *Dropout rate:* 12 months: 20% (CC) and 13% (UC) and 24 months: 47% (CC) and 43% (UC) *Effect size differences between CC and UC in mean rating scale scores in all highly anxious patients at 12 and 24 month follow-up*[a] Effect size (95% CI); *P*-value at 12 months (n=208)[b] SF-36 MCS: 0.38 (0.13; 0.63); *P* = 0.003 SIGH-A: 0.30 (0.05; 0.55); *P* = 0.02 PDSS[c]: 0.42 (0.08; 0.76); *P* = 0.02 GADSS[d]: 0.41 (0.15; 0.66); *P* = 0.002 SF-36 PCS: 0.06 (-0.19; 0.31); *P* =0.62 PHQ-9[e]: 0.45 (0.19; 0.71); *P* = 0.001 Effect size (95% CI); *P*-value at 24 months (n=138)[b] SF-36 MCS: 0.17 (-0.07; 0.42); *P* = 0.17 SIGH-A: 0.30 (0.05; 0.54); *P* = 0.02 PDSS[c]: 0.19 (-0.15; 0.53); *P* = 27 GADSS[d]: 0.06 (-0.19; 0.32); *P* = 0.63 SF-36 PCS: 0.01 (-0.24; 0.25); *P* =0.96 PHQ-9[e]: 0.22 (-0.04; 0.48); *P* = 0.10 [a] Positive scores indicate improvement due to the intervention (CC) vs. UC at the given time point [b] Of the 250 highly anxious patients randomized at baseline [c] PD and PD/GAD cohorts only [d] GAD and PD/GAD cohorts only [e] Major depression cohort only *Proportion of patients*[a] *achieving ≥ 40% decline from baseline levels of anxiety or mood symptoms at 12 months follow-up* SIGH-A, n/N (%) CC: 52/99 (53) vs. UC: 35/108 (32); *P* = 0.003; NNT (95% CI): 5 (14; 3) PDSS[b], n/N (%) CC: 35/50 (70) vs. UC: 21/55 (38); *P* = 0.001; NNT (95% CI): 3 (7;2) GADSS[c], n/N (%) CC: 48/97 (49) vs. UC: 28/100 (28); *P* = 0.002; NNT (95% CI): 12;3) PHQ-9[d], n/N (%) CC: 56/87 (64) vs. UC: 40/98 (41); *P* = 0.001; NNT (95% CI): 4 (10; 3) [a] Of the 250 highly anxious patients randomized at baseline [b] PD and PD/GAD cohorts only [c] GAD and PD/GAD cohorts only	The authors concluded that *"Telephone-delivered, centralized, stepped CC improves HRQoL, anxiety, and mood symptoms. These improvements were durable and particularly evident among those most anxious at baseline, and among African-Americans and men."*[24] p245

Main Study Findings	Authors' Conclusions
[a] Major depression cohort only *Safety outcomes:* none reported.	
Depression and Anxiety	
van den Berg, 2015[25]	
Baseline characteristics: -Total study population (N=123); mean (SD) age: 44.0 years (12.5); females: 71.5% -91% had at least 1 diagnosis in the spectrum of mood disorders *Dropout rate:* 8.1% (Group 1 [2.4%], Group 2 [7.5%], Control [14.6%]) Of all patients with follow-up, 62 (55%) had at least 1 contact with a psychotherapist between baseline and follow-up *Primary outcome: Change in BSI-18 symptom scores at 6 months, mean (SD) (ITT)* Results are for Anxiety/Depression/Somatization Baseline: 7.60 (5.09)/8.71 (5.22)/4.73 (4.07) Group 1: 6.71 (5.69)/6.27 (5.75)/4.76 (3.90) Group 2: 5.38 (4.02)/6.22 (5.59)/4.70 (3.99) Control group: 6.37 (5.80)/6.06 (5.70)/3.91 (4.31) Note: Group 1 = regular telephone contacts; Group 2: regular telephone contacts + text messaging; Control: usual care; Analysis was by ANCOVA adjusting for baseline values of the corresponding BSI-18 score The average BSI-18 anxiety score after 6 months was -2.04 points lower in Intervention 2 vs. control (*P*=0.042). The difference in depression score had an average treatment effect of -1.73 (*P*=0.097). In an exploratory sensitivity analysis restricted to 75% of patients with the highest symptom scores at baseline, Intervention 1 had a significant effect for both anxiety (*P*=0.036) and depression (*P*=0.046). *Safety outcomes:* none were reported.	The authors concluded that *"Telemedicine provides a novel option in psychiatric ambulatory care with statistically significant effects on anxiety. A positive tendency was observed for depression, especially in cases with higher symptom load at baseline."*[25] p82
Post-Traumatic Stress Disorder	
Franklin, 2017[26]	
Baseline characteristics: -Total study population (N=18); Full PTSD: 67%; subthreshold PTSD: 33%; mean (SD) age: 53.8 years (12.0); Male: 100%; African-American: 66% *Dropout rate (3-month follow-up):* 33.3% (5 of 11 [45%] CBT-T and 1 of 7 [14%] in CBT-IP) *Mean (SD) PSQI response by assessment point*[a] CBT-IP Baseline (n=7): 15.1 (3.3) Post-treatment (n=6): 11.3 (4.0) 1 month (n=6): 9.8 (4.9) 3 months (n=6): 11.8 (5.5) CBT-T Baseline (n=10): 14.1 (2.7) Post-treatment (n=7): 12.6 (5.2) 1 month (n=7): 12.0 (4.4) 3 month (n=6): 10.5 (5.5) Both groups combined Baseline (n=17): 14.5 (2.9)	The authors concluded that *"Specifically, current pilot data suggest that telephone-delivered CBT-I may be able to reduce trauma-related insomnia symptoms. Future trials are needed to assess the effectiveness of CBT-I delivered to rural veterans with*

Main Study Findings	Authors' Conclusions
Post-treatment (n=13): 12.0 (4.5) 1 month (n=13): 11.0 (4.5) 3 month (n=12): 11.0 (5.3) [a] Descriptive statistics only A hierarchical longitudinal linear regression analysis, with PSQI scores nested within patients, was conducted to examine PSQI scores over time. The "group x time" interaction was not significant (*P*=0.18) indicating that the groups did not differ over time in their PSQI scores. The effectiveness of telephone-delivered CBT-I using between and within-group effect sizes was done using Cohen's *d*. Large between-group effect sizes favoured in-person CBT-I both immediately (*d*=0.77) and at one month (*d*=1.06), but was less at three-months (*d*=0.10). For within-group effect sizes, large effects for in-person CBT-I at all three time points was found (ds range from 1.00 to1.61) and medium to large effects for telephone-delivered CBT-I (ds range from .55 to 1.33). Patient acceptability was 83% across both groups *Safety outcomes:* none reported	*post-traumatic insomnia."*[26] p1

Adj = adjusted; AE = adverse events; AIDS = acquired immune deficiency syndrome; ANCOVA = analysis of covariance; BATD = behavioral activation treatment for depression; BDI-II: Beck Depression Inventory II; BHS = Beck Hopelessness Scale; BSI-18 = Brief Symptom Inventory Short Version; CBT = cognitive behavioral therapy; CBT-IP = CBT in person; CBT-T = CBT by telephone; CC = collaborative care; CGI-I: Clinical Global Impressions - Improvement Questionnaire; CGI-S = Clinical Global Impressions - Severity of Illness Questionnaire; COPD = chronic obstructive pulmonary disease; CSQ = Client Satisfaction Questionnaire; DSM -IV = Diagnostic and Statistical Manual (of Mental Disorders) IV; EQ-5D-5L = EuroQoL 5 Dimension 5 Level; GAD = generalized anxiety disorder; GADS-7 =GAD 7-item Scale; GADSS = Generalized Anxiety Disorder Severity Scale; GDS = Geriatric Depression Scale; HAMD-17 = 17-item Hamilton Depression Rating Scale; HIV = human immunodeficiency virus; HRQoL = health-related quality of life; IASMHS = Inventory of Attitudes Tow ard Seeking Mental Health Services; ITT = intention-to-treat; MDD = major depressive disorder; MMRM = linear mixed effect repeated measures; NC = not calculated; Q-LES-Q: Chinese-translated version of the 'short-form' of the Quality of Life Enjoyment and Satisfaction Questionnaire; PD = panic disorder; PDSS = Panic Disorder Severity Scale; PGI = physician global impression; PHQ-9 = Patient Health Questionnaire 9; PRIME-MD = Primary Care Evaluation of Mental Disorders Anxiety Module; PSQI = Pittsburgh Sleep Quality Index; PTSD = post-traumatic stress disorder; RCT = randomized controlled trial; SCID = Structured Clinical Interview for the DSM-IV Axis I Disorders; SCL-20 = Symptom Checklist -20 Depression Scale; SD = standard deviation; SF-36 MCS = Medical Outcomes Study Short-Form 36 Mental Component Summary; SIGN-A = Hamilton Anxiety Rating Scale; SMD = standardized mean difference; TAU = treatment as usual; TBI = traumatic brain injury; T-CSCT = telepsychiatry-based culturally sensitive collaborative treatment; T-IPT = interpersonal psychotherapy by telephone; UC = usual care; Unadj = unadjusted; VC = videoconferencing