ABSTRACT

Introduction Major advancements in technology have led to considerations how telemedicine (TM) and other technology platforms can be meaningfully integrated in treatment for psychiatric disorders. The COVID-19 pandemic has placed a further focus on use of TM in psychiatry. Despite the widespread use of TM, little is known about its effect compared with traditional in-person (IP) consultation. The objective of this systematic review is to examine if individual psychiatric outpatient interventions for adults using TM are comparable to IP in terms of (1) psychopathology outcomes, (2) levels of patient satisfaction, (3) working alliance and (4) dropout from treatment.

Methods and analysis This review will only include randomised controlled trials for adult participants with mood disorders, anxiety or personality disorders. The primary outcome is psychopathology, and secondary outcomes include patient satisfaction, treatment alliance and dropout rate. Systematic searches were conducted in MEDLINE, APA PsycINFO, Embase, Web of Science and CINAHL. The inverse-variance method will be used to conduct the meta-analysis. Effect sizes will be calculated as standardised mean difference (Hedges' g) for the primary outcome, mean difference for patient satisfaction and working alliance, and risk ratio for the dropout rate. Effect sizes will be supplemented with 95% CI. We will calculate the F statistic to quantify heterogeneity and Chi-square statistic (χ²) to test for heterogeneity for the primary outcome. Potential clinical and methodological heterogeneity moderators will be assessed in subgroup and sensitivity analysis. The risk of bias will be assessed by Cochrane Risk of Bias Tool V.2, and confidence in cumulative evidence will be assessed by Grading of Recommendations Assessment, Development and Evaluation.

Ethics and dissemination No ethical approval is required for this systematic review protocol. Data sets will be deposited in the Zenodo repository. The findings of this study will be published in a peer-review scientific journal.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This review will include randomised controlled trials to compare individual psychiatric treatment using telemedicine (TM) or in-person (IP) for people with mood, anxiety or personality disorders.

⇒ Validated and standardised measures will be used to assess psychopathology, patient satisfaction (Client Satisfaction Questionnaire-8) and working alliance (Working Alliance Inventory) across all the studies.

⇒ This systematic review will calculate and compare dropout rates between TM and IP treatment formats. An outcome that has not previously been examined in a systematic review.

⇒ The stringent eligibility criteria regarding study design, participants, interventions and outcome measures will result in some studies being excluded.

INTRODUCTION

Telemedicine (TM) is, according to the WHO, interpreted as ‘healing at a distance’ that enables remotely delivered treatment while the patient and clinician are in a different physical location. Various names have been suggested and used interchangeably in the scientific literature to describe TM. Telehealth, telepsychiatry, video consultation, video conference, telemental health and teleconsultation, for example, are commonly used. TM is the broader term and covers synchronous (video, telephone) and asynchronous (‘store and forward’, ie, emails, SMS) technologies.2

Experimentation with TM in medical settings first began in the 1950s. These studies were based on a simple two-way closed-circuit television and TM was used for treatment and education purposes. Advances in technology and increasing access to the internet mean that TM can now be quickly accessed using a smartphone or other digital devices.
Furthermore, the COVID-19 pandemic has led to the accelerated use of digital solutions in healthcare systems in many countries.2,3 Currently, there are a number of large ongoing trials comparing TM to in-person (IP) in populations of depression, anxiety, obsessive–compulsive disorder and perinatal women.10–12

The use of TM in mental health services has several potential advantages such as making psychiatric services more accessible and flexible, reducing the cost of transport and time, reducing stigma, promoting patient autonomy and providing an opportunity for people with mental health difficulties to engage with services if they find it challenging to attend IP consultations.13–14

There are also some potential disadvantages of TM, which include concerns about data security, technical obstacles, questions regarding the efficacy of interventions grounded in TM, which patient groups TM is most suitable for, concerns about establishing good working alliances, maintaining treatment engagement and the allocation of resources of trained clinicians.15–16 Different populations (eg, geriatric, suicidal or perinatal) can also experience a range of barriers and challenges using TM such as issues of privacy and safety, difficulty learning new technologies or the provision of care for acute mental health problems.17–19

Over the last two decades, several systematic reviews have compared TM with IP within psychiatry.20 21 These systematic reviews indicate that TM for psychiatric outpatients is equivalent to IP consultations regarding efficacy (psychopathology, patient satisfaction and working alliance). Unfortunately, the majority of these reviews have usually been descriptive in nature and included trials of varying quality. Furthermore, many of the randomised controlled trials (RCTs) included in these reviews have been underpowered. Currently, there is a lack of quantitative analyses to determine the efficacy of psychiatric treatment provided by TM compared with IP formats.

Three meta-analyses conducted by Drago et al, Batastini et al and Giovanetti et al have examined outcomes comparing TM to IP treatment.22–24 Drago et al examined a wide range of interventions within psychiatry but excluded psychotherapeutic interventions. They found that TM was not inferior to IP across a range of mental health outcomes. Batastini et al carried out a large review of TM and IP for a broad range of psychotherapeutic interventions within mental health and they found no significant differences in outcomes between the two treatment formats. Batastini et al’s review included a range of study designs (randomised and non-randomised trials) and different treatment formats (individual and group) across a broad range of mental health related outcomes (symptoms, hospitalisation, relapse, medication compliance). Giovanetti et al conducted a systematic review and meta-analysis to examine the treatment effect for patients with depressive symptoms. Their meta-analysis included 11 RCT studies that directly compared individual psychotherapy through TM with IP and they found no significant differences in outcomes between the two treatment formats. Results from the three reviews conducted indicate that treatment using TM is comparable to IP treatment, although the three research groups also acknowledge a number of limitations with their respective reviews. They recommend that further trials and reviews were necessary and highlight the need for more rigorous study designs, inclusion of a broader range of psychiatric disorders, clearly defined interventions and detailed diagnostic descriptions to develop the evidence base when comparing TM and IP interventions.

Satisfaction outcomes in studies comparing TM to IP in psychiatric outpatients have been assessed in a single meta-analysis by Hyler et al.25 This review concluded that there were no differences in levels of patient satisfaction between TM and IP modalities although the authors noted that only a few studies used standardised satisfaction instruments. A number of studies applied ad hoc or untested satisfaction instruments, where the reliability or validity was not reported. It is essential to use standardised and empirically evaluated measures to allow meaningful comparisons between different studies.26

Working alliance was assessed in meta-analysis conducted by Norwood et al that concluded that alliance in TM treatment was inferior to IP treatment.27 This finding contrasts with other systematic reviews that suggest that alliance in individual treatment using TM was equal or better than IP treatment.21 28 29

Currently, there is no meta-analysis on dropout rates in treatment using TM compared and IP making it a research area that needs to be addressed.

Based on the current research examining interventions using TM compared with IP consultations, there is a need to conduct a meta-analysis covering a range of psychiatric disorders and focusing on multiple clinical outcomes. This meta-analysis will build on previous research and address some of the current limitations in the literature by conducting a systematic review including studies with rigorous study design (only RCT’s), defined clinical interventions (individual treatment), specific psychiatric populations (diagnoses of anxiety, depression or personality disorder) using standardised assessments for psychopathology, working alliance and treatment satisfaction.

The specific objective of this systematic review is to examine if individual psychiatric outpatient interventions for adults conducted using TM are comparable to IP in terms of (1) psychopathology outcomes, (2) levels of patient satisfaction, (3) working alliance and (4) dropout from treatment.

**METHODS AND ANALYSIS**

This protocol will be conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis Protocol (PRISMA-P).30 The PRISMA-P checklist can be found in the online supplemental file 1. The review has been registered in the PROSPERO International Prospective Register of Systematic Reviews (Registration number
Trials using asynchronous communications systems as ► Trials involving populations that primarily treating ► Different psychotherapy (treatments) approaches in Group therapy. ► Participant<18 years. ► Treatment is defined as intervention involving psychotherapy, pharmacological treatment or psychoeducation.

Types of intervention
Individual treatment through synchronous real-time video delivered sessions/consultations in outpatient settings.

Types of comparator/control
Individual treatment IP and same active treatment as the intervention group receives.

Types of outcomes
Eligible studies have assessed psychopathology following an intervention (eg, mails and static website without video function) and telephone (only audio) as the intervention will not be included.

Exclusion criteria
► Non-RCT studies. ► Participant<18 years. ► Different psychotherapy (treatments) approaches in intervention and control group. ► Trials involving populations that primarily treating psychotic disorders, mental retardation, bipolar disorders, alcohol abuse and substance use disorders will be excluded. ► Trials using asynchronous communications systems as an intervention (eg, mails and static website without video function) and telephone (only audio) as the intervention will not be included.

Information sources and search strategy
The first step in the systematic review has been a comprehensive search in electronic databases. The database search strings were created in January 2021 by AAS with guidance from the information specialist Trine Kestel, who has expertise in systematic review searching (psychiatric research unit, Region Zealand). The database search strategy was developed with input from the project team. Search strategies are provided in the online supplemental file 2.

The databases used for the searches are as follow: Medline (PubMed interface, 1986 onwards), APA PsycINFO (OVID interface, 1967 onwards), Embase (OVID interface, 1974 onwards), Web of Science (Clarivate interface, 2001 onwards) and CINAHL (EBSCOhost interface, 1981 onwards).

Medical subject headings and text words related to the search terms ‘psychiatry’ and ‘telemedicine’ were used for developing the search string in MEDLINE. Both search terms—psychiatry and telemedicine—were then combined with [AND]. Specific syntax and subject headings were subsequently adapted individually to the different databases.

No language and date restriction was implemented in the search process. Due to the unmanageable results (>20 000 hits) in the preliminary search, Cochrane’s highly sensitive search strategy filters identifying randomised trials has been applied in the final search string. Repeated search will be performed prior to the final analysis to identify further eligible studies. Unpublished studies will not be sought.

The second step in the search strategy will be a manual literature search to identify additional primary studies for the systematic review.

The third step will be scanning the reference lists of included studies or relevant reviews identified in the first and second steps.

Data management
Records from the literature search will be exported to the reference manager Endnote V.X9. From Endnote records will be exported to Covidence. Covidence is a web-application tool that facilitates collaboration among the review team members during the study selection and data extraction process. Extracted data in Covidence will be exported to RevMan V.5.4 for data analysis.

Selection process
AAS and SFA will be responsible for the selection process. In the first step, the two authors will independently screen the title and abstracts of the records in Covidence to identify potentially eligible records. The Second step will be obtaining and screening full-text reports to decide if reports meet eligible criteria. Disagreement through the selection process will be resolved by discussion between the two authors. In case of continued disagreement despite discussion, a third reviewer, OFS will be consulted. The selection process—including exclusion reasons—will be documented in the PRISMA-P flow diagram. Inter-rater reliability will be measured by Cohen’s kappa coefficient (κ) for the (1) title and abstract screening process and (2) full-text review process.
Data collection process
AAS and SFA will be responsible for the data collection process. Data extraction will be carried out through a standardised electronic data extraction form in Covidence. The data extraction form will initially be piloted on some reports, and the reviewers will meet and discuss the form before starting the review. Disagreement through the data collection process will be resolved by discussion between the two authors. In case of continued disagreement despite discussion, a third reviewer (OJS) will be consulted. If we encounter multiple reports of the same study, we will extract data from all reports into a single data collection form in Covidence.34 Missing data will be obtained by contacting and requesting these data from the study authors.

Data items
We will extract the following data items for each study: (a) study characteristics (authors, author contact details, aim of the study, trial design, location, trial size, sample size calculation, year of publication and country), (b) population characteristics (remote/rural area or urban, country, diagnosis/condition, comorbidity, mean age and gender), (c) intervention/control (internet connection speed, bandwidth, therapy type, number of consultation sessions and duration of consultation) and (d) clinical outcome (assessments tool, psychopathology, satisfaction, alliance, and dropout rate). When reported in the studies, we will collect data from the ‘intention-to-treat’ analysis; otherwise, perprotocol data will be collected. For crossover RCTs, only data before crossover will be used to prevent carryover effects and units of analysis errors.

Outcomes and prioritisation
The primary outcome in this review is psychopathology assessed by clinician or patient-rated scales. As we expect that different assessments tools have been used for measuring the primary outcome, we will prioritize clinician-rated scales and secondary patient-rated scales.

The secondary outcomes in the review will be (a) patient satisfaction, (b) working alliance and (c) dropout rate. Satisfaction must be assessed by Client Satisfaction Questionnaire-8 (CSQ-8),26 and the working alliance must be assessed by the client Working Alliance Inventory (WAI)35 in the included studies. The dropout rate is defined as the proportion of individuals who withdrew after being randomised to the total number of participants randomised to a condition.

Risk of bias in individual studies
AAS and SFA will perform the risk-of-bias (quality) assessment in the individual studies. The revised Cochrane Risk of Bias Tool V.2 for randomised trials34 will be applied. Our primary outcome—psychopathology—will be assessed for risk of bias in each study. The bias domain that will be assessed include (a) bias arising from the randomisation process, (b) bias due to deviations from intended interventions, (c) bias due to missing outcome data, (d) bias in the measurement of the outcome and (e) bias in the selection of the reported result. Overall risks of bias for each study outcome will be marked as: (1) ‘low risk of bias’ if all domains are judged to be low, (2) ‘some concerns’ if at least one domain are judged to raise some concerns but not to be at high risk of bias for any domain and (3) ‘high risk of bias’ if any domain is judged to be at high risk of bias. Disagreement between the mentioned researchers regarding the risk of bias will be resolved through consensus or a third researcher (OJS). Covidence tool will be used to assess the risk of bias.

Data synthesis
The general strategy for our data synthesis is to perform a quantitative synthesis (meta-analysis). A narrative synthesis will be performed if heterogeneity (I²) is substantially high and will include summary tables and descriptions of the findings. I² values will be judged as follows: 0%–40% may represent little heterogeneity, 30%–60% may represent moderate heterogeneity, 50%–90% may represent substantial heterogeneity and 75%–100% represent considerable heterogeneity. Heterogeneity, which is the percentage of variation across studies that is due to heterogeneity rather than chance, will be grouped in (1) clinical, (2) methodological and (3) statistical heterogeneity.34 Clinical heterogeneity refers to the variation across studies regarding age, sex, diagnosis, treatment site, and intervention characteristics (duration of intervention, number of interventions and time interval between interventions). Methodological heterogeneity refers to the variability in the risk of bias and outcome measurement tools. Statistical heterogeneity refers to the differences in the intervention effects of each trial being evaluated.

Quantitative synthesis
We will use the inverse-variance method for carrying out the meta-analysis. Larger studies with less variance will be given more weight in the meta-analysis due to more precise effect size estimates than smaller studies. As we expect clinical and methodological heterogeneity in the pooled studies, we will use the random-effects model to obtain the overall effect size estimate. When heterogeneity is low a fixed-effect model will be chosen.

Continuous outcome measures
We will calculate the standardised mean difference (SMD) effect size for the primary outcome using Hedges’ g formula. Because we intend to use different assessments tools to calculate the effect size for the primary outcome in each study, SMD will be statistically suitable for this. Forest plot will be used for presenting effect sizes and overall effect size. A 95% CI will supplement the calculated effect sizes. Furthermore, we will calculate the I² statistic to quantify heterogeneity and Chi-square statistic (χ²) to test for heterogeneity (p≤0.1 significance level).

For the secondary outcomes—satisfaction and working alliance—we will calculate mean difference effect size
as we have restricted these secondary outcomes to be assessed by a standardised tool (CSQ-8 and WAI). Therefore, standardising is not needed to calculate the effect size across the studies. Beyond this, the same statistical approach for the primary outcome already described will be applied to the secondary outcomes satisfaction and working alliance.

We intend to combine ‘end of treatment’ scores (post intervention) and ‘change score’ data (changes from baseline) to calculate the estimated overall effect size for both primary and secondary outcomes. This is a valid approach. If the change score is not reported or cannot be calculated, postintervention data will be used as the second choice.

**Dichotomous outcome measures**

We will calculate the risk ratio effect size and its 95% CI for the secondary outcome (dropout rate). Forest plot will be used to present effect sizes and overall effect size and supplemented with I² and χ² statistics. We define dropout as the number of participants not completing scheduled treatment courses, that is, the difference in the number of participants who started the first treatment session (baseline) and completed the treatment course (post-treatment).

**Additional primary outcome analyses (investigating heterogeneity)**

For the primary outcome, a subgroup analysis (a) for different patient groups will be performed based on participant diagnosis as specified in the eligibility criteria, (b) sex, (c) ages, (d) length of treatment course / programme, (e) therapy type, (f) settings (remote / rural area or urban) and (g) vulnerable populations (eg, perinatal, ethnically/racially diverse and geriatric populations).

A sensitivity analysis will be performed to determine the robustness of the meta-analysis and will include: 1. Removing low-quality studies and repeating the meta-analysis. 2. Testing for any possible difference between ‘end-of-treatment’ scores and ‘change scores’. 3. Testing for whether the findings are sensitive to random effects or fixed effects models. 4. Assessing the effect of the year of publication; a meta-regression will be performed, and a p value for the regression will be calculated (p<0.05 significance level). The rationality for this meta-regression is to analyse if the technological or therapeutic evolution affects the primary outcome.

**Meta bias**

Publication bias will be assessed and will be done by visually assessing a funnel plot supplied by Egger’s test. If the funnel plot is not symmetrical, the publication bias is present. This is investigated by using the Begg’s rank correlation test and the Egger’s linear regression test. A funnel plot is used to determine whether there is any asymmetry which may suggest publication bias. The funnel plot is constructed from a meta-analysis of studies grouped by the size of the study, with smaller studies placed at the bottom of the plot and larger studies placed at the top of the plot. The horizontal axis represents the effect size, while the vertical axis represents the number of studies. The funnel plot is used to assess the symmetry of the data, with the expectation that the studies on the left and right sides of the plot should be similar in size. If the funnel plot is asymmetrical, this may indicate that smaller studies with smaller effect sizes are missing, which could be due to publication bias.

**Confidence in cumulative evidence**

We will use the Grading of Recommendations Assessment, Development and Evaluation approach as recommended by Cochrane Collaboration to assess the confidence of the body of evidence.
Open access

12 Singla DR, Meltzer-Brodsky SE, Silver RK, et al. Scaling up maternal mental healthcare by increasing access to treatment (summit) through non-specialist providers and telemedicine: a study protocol for a non-inferiority randomized controlled trial. Trials 2021;22.
32 Covidence systematic review software, veritas health innovation, Melbourne, Australia. Available: www.covidence.org
37 Egger M, Smith GD, Schneider M. Papers bias in meta-analysis detected by a simple, graphical test.