Evidence Brief: Safety and Effectiveness of Telehealth-delivered Mental Health Care

Supplemental Materials

October 2022

U.S. Department of Veterans Affairs
Veterans Health Administration
Health Services Research & Development Service

Recommended citation: Beech EH, Young S, Anderson JK, Belsher BE, Parr NJ. Evidence Brief: Safety and Effectiveness of Telehealth-delivered Mental Health Care. Washington, DC: Evidence Synthesis Program, Health Services Research and Development Service, Office of Research and Development, Department of Veterans Affairs. VA ESP Project #09-199; 2022.

TABLE OF CONTENTS

Appendix A: Search Strategy 1
Systematic Reviews 1
Primary Studies 6
Appendix B: Hand-searched Systematic Reviews 8
Appendix C: Excluded Studies 10
Appendix D: Evidence Tables 13
Characteristics of Included Primary Studies 13
RCTs 13
Cohort Studies 22
Outcome Data of Included Primary Studies 28
RCTs 28
Cohort Studies 58
Quality Assessment of Included Primary Studies 78
ROB 2 78
ROBINS-I 105
Strength of Evidence for Included Studies 127
Strength of Evidence for KQ 1 127
Strength of Evidence for KQ 2 136
Appendix E: Peer Review Disposition 138
References 143

APPENDIX A: SEARCH STRATEGY

SYSTEMATIC REVIEWS

1. Search for current systematic reviews (limited to last 7 years) **Date Searched: 05-25-22**			
A. Bibliographic Databases:	**#**	**Search Statement**	**Results**
MEDLINE: Systematic Reviews Ovid MEDLINE(R) ALL 1946 to May 24, 2022	1	exp Stress Disorders, Post-Traumatic/ or Depression/ or exp Depressive Disorder/ or exp Anxiety/ or Bipolar Disorder/ or exp Substance-Related Disorders/ or exp Suicide/ or exp Schizophrenia/ or (post?traumatic stress disorder or post?traumatic neuros* or PTSD or moral injury or depression or depressive or melancholia or anxiety or angst or anxieties or hypervigilance or (bipolar adj1 disorder*) or bipolar affective disorder or bipolar affective psychosis or manic?depressive psychosis or manic disorder or substance?related disorder or ((drug or substance) adj1 disorder) or ((drug or substance or chemical) adj1 addiction) or suicid* or serious mental illness or SMI or schizophreni*).ti,ab.	1143016
	2	exp Telemedicine/ or exp Videoconferencing/ or Electronic Mail/ or exp Telephone/ or Text Messaging/ or exp Mobile Applications/ or exp Cell Phone/ or (telemedicine or telehealth or mHealth or eHealth or (mobile adj1 health) or video?conferenc* or e?mail or online portal or telephone or cell phone or mobile phone or text message or SMS or mobile app*).ti,ab.	162994
	3	(teen* or adolescen* or child* or infant* or youth).ti,ab.	2053569
	4	1 AND 2	14937
	5	4 NOT 3	12665
	6	(systematic review.ti. or meta-analysis.pt. or meta-analysis.ti. or systematic literature review.ti. or this systematic review.tw. or pooling project.tw. or (systematic review.ti,ab. and review.pt.) or meta synthesis.ti. or meta-analy*.ti. or integrative review.tw. or integrative research review.tw. or rapid review.tw. or umbrella review.tw. or consensus development conference.pt. or practice guideline.pt. or drug class reviews.ti. or cochrane database syst rev.jn. or acp journal club.jn. or health technol assess.jn. or evid rep technol assess summ.jn. or jbi database system rev implement rep.jn. or (clinical guideline and management).tw. or ((evidence based.ti. or evidence-based medicine/ or best practice*.ti. or evidence synthesis.ti,ab.) and (((review.pt. or diseases category/ or behavior.mp.) and behavior mechanisms/) or therapeutics/ or evaluation studies.pt. or validation studies.pt. or guideline.pt. or pmcbook.mp.)) or (((systematic or systematically).tw. or critical.ti,ab. or study selection.tw. or ((predetermined or inclusion) and criteri*).tw. or exclusion criteri*.tw. or main outcome measures.tw. or standard of care.tw. or standards of care.tw.) and ((survey or surveys).ti,ab. or overview*.tw. or review.ti,ab. or reviews.ti,ab. or search*.tw. or handsearch.tw. or analysis.ti. or critique.ti,ab. or appraisal.tw. or (reduction.tw. and (risk/ or risk.tw.) and (death or recurrence).mp.)) and ((literature or articles or publications or publication or bibliography or bibliographies or published).ti,ab. or pooled data.tw. or unpublished.tw. or citation.tw. or citations.tw. or database.ti,ab. or internet.ti,ab. or textbooks.ti,ab. or references.tw. or scales.tw. or papers.tw. or datasets.tw. or	518675

		trials.ti,ab. or meta-analy*.tw. or (clinical and studies).ti,ab. or treatment outcome/ or treatment outcome.tw. or pmcbook.mp.))) not (letter or newspaper article).pt.	
	7	5 AND 6	829
	8	limit 7 to english language	687
CDSR: Protocols and Reviews EBM Reviews - Cochrane Database of Systematic Reviews 2005 to May 18, 2022	1	(Post-Traumatic Stress Disorder or Depression or Depressive Disorder or Anxiety or Bipolar Disorder or Substance-Related Disorders or Suicide or Schizophrenia).kw. or (post?traumatic stress disorder or post?traumatic neuros* or PTSD or moral injury or depression or depressive or melancholia or anxiety or angst or anxieties or hypervigilance or (bipolar adj1 disorder*) or bipolar affective disorder or bipolar affective psychosis or manic?depressive psychosis or manic disorder or substance?related disorder or ((drug or substance) adj1 disorder) or ((drug or substance or chemical) adj1 addiction) or suicid* or serious mental illness or SMI or schizophreni*).ti,ab.	1098
	2	(Telemedicine or Videoconferencing or Electronic Mail or Telephone or Text Messaging or Mobile Applications or Cell Phone).kw. or (telemedicine or telehealth or mHealth or eHealth or (mobile adj1 health) or video?conferenc* or e?mail or online portal or telephone or cell phone or mobile phone or text message or SMS or mobile app*).ti,ab.	168
	3	(teen* or adolescen* or child* or infant* or youth).ti,ab.	3003
	4	1 AND 2	34
	5	4 NOT 3	27

1. Search for current systematic reviews (limited to last 7 years) Date Searched: 05-25-22		
B. Non-bibliographic databases	**Evidence**	**Results**
AHRQ: evidence reports, technology assessments, U.S Preventative Services Task Force Evidence Synthesis	http://www.ahrq.gov/research/findings/evidence-based-reports/search.html Search: mental health; telehealth Agarwal S, Jalan M, Wilcox HC, Sharma R, Hill R, Pantalone E, Thrul J, Rainey JC, Robinson KA. Evaluation of Mental Health Mobile Applications. Technical Brief 41. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 75Q80120D00003.) AHRQ Publication No. 22-EHC016. Rockville, MD: Agency for Health care Research and Quality; May 2022. DOI: https://doi.org/10.23970/AHRQEPCTB41. Technical Brief: Telehealth: Mapping the Evidence for Patient Outcomes From Systematic Reviews. Content last reviewed January 2020. Effective Health Care Program, Agency for Health care Research and Quality, Rockville, MD. https://effectivehealth care.ahrq.gov/products/telehealth/technical-brief	2
CADTH	https://www.cadth.ca CADTH. Post-Traumatic Stress Disorder: Summary of Evidence of the Clinical Effectiveness of Treatments. 2018.	11

	CADTH. Tele-medicine for Patients with Mental Health Disorders: Clinical and Cost-effectiveness. 2015. CADTH. Internet-Based Cognitive Behavioral Therapy for Post-Traumatic Stress Disorder: A Review of Clinical Effectiveness. 2018. CADTH. Internet-Delivered Cognitive Behavioural Therapy for Major Depressive Disorder and Anxiety Disorders: A Health Technology Assessment. 2019. CADTH. Telehealth Services for the Treatment of Psychiatric Issues: Clinical Effectiveness, Safety, And Guidelines. 2015. CADTH. Telehealth for the Assessment and Treatment of Depression, Post-Traumatic Stress Disorder, and Anxiety: Clinical Evidence. 2018. CADTH. Telehealth-Delivered Opioid Agonist Therapy for the Treatment of Adults with Opioid Use Disorder: Review of Clinical Effectiveness, Cost-Effectiveness, and Guidelines. 2018. CADTH. Telehealth and Mobile Services for Substance Use Disorder: Clinical Effectiveness, Cost-Effectiveness and Guidelines. 2020. CADTH. Virtual Health Care for Adults with Schizophrenia and/or Psychosis: Clinical Effectiveness and Guidelines. 2020. CADTH. e-Therapy Interventions for the Treatments of Patients with Depression: A Review of Clinical Effectiveness. 2018. CADTH. e-Therapy Interventions for the Treatments of Substance Use Disorders and Other Addictions: A Review of Clinical Effectiveness. 2018. CADTH. e-Therapy Interventions for the Treatment of Post-Traumatic Stress Disorder: Clinical Evidence. 2018. CADTH. Internet-Based Cognitive Behavioral Therapy for Post-Traumatic Stress Disorder: A Review of Clinical Effectiveness. 2018. CADTH. e-Therapy Interventions for the Treatments of Patients with Depression: A Review of Clinical Effectiveness. 2018. CADTH. e-Therapy Interventions for the Treatment of Anxiety: Clinical Evidence. 2018. CADTH. Telehealth Services for the Treatment of Psychiatric Conditions: Clinical Effectiveness, Safety, and Guidelines. 2015.	
ECRI Institute	https://guidelines.ecri.org/ VA/DoD clinical practice guideline for the management of substance use disorders. 2021.	1
HTA: Health Technology Assessments (UP TO 2016)	http://www.ohsu.edu/xd/education/library/ See CDSR search above	0
EPPI-Centre	http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=62	0

	Use browser search function [CNTL + F] for keyword search	
NLM	http://www.ncbi.nlm.nih.gov/books Salisbury C, O'Cathain A, Thomas C, et al. An evidence-based approach to the use of telehealth in long-term health conditions: development of an intervention and evaluation through pragmatic randomised controlled trials in patients with depression or raised cardiovascular risk. Southampton (UK): NIHR Journals Library; 2017 Jan. (Programme Grants for Applied Research, No. 5.1.) Center for Substance Abuse Treatment (US). Using Technology-Based Therapeutic Tools in Behavioral Health Services. Rockville (MD): Substance Abuse and Mental Health Services Administration (US); 2015. (Treatment Improvement Protocol (TIP) Series, No. 60.) Veazie S, Bourne D, Peterson K, et al. Evidence Brief: Video Telehealth for Primary Care and Mental Health Services. Washington (DC): Department of Veterans Affairs (US); 2019 Feb.	3
VA Products - VATAP, PBM and HSR&D publications	A. http://www.hsrd.research.va.gov/research/default.cfm B. http://www.research.va.gov/research_topics/ C. https://va.dimensions.ai/discover/publication Telehealth Cognitive-Behavioral Therapy for Depression in Parkinson's Disease Alejandro Interian PhD Lyons Campus of the VA New Jersey Health Care System, Lyons, NJ Lyons, NJ Funding Period: July 2016 - September 2020 Impact of Combined Recovery Program and Home Telehealth Among Veterans with substance use disorders in the VA Inpatient Setting Elizabeth J. Santa Ana PhD MA BA Charleston, SC Funding Period: October 2021 - September 2025 Telehealth CBT to increase engagement in pain treatment among Veterans using prescription opioids Lisham Ashrafioun PhD Canandaigua, NY Funding Period: May 2022 - April 2026	3

2. Search for systematic reviews currently under development (includes forthcoming reviews & protocols) **Date Searched: 05-25-22**		
D. Under development:	**Evidence:**	**Results:**
AHRQ topics in development (EPC Status Report)	Email Charli Armstrong Charlotte.Armstrong1@va.gov	0

PROSPERO (SR registry)	http://www.crd.york.ac.uk/PROSPERO/ Cristiane de Cássia Bergamaschi, Reginaldo Tavares Franquez. Effectiveness of e-health technologies for treatment of depression, anxiety and emotional distress in people with diabetes mellitus: systematic review. PROSPERO 2022 CRD42022314773 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022314773 Adam Lewkowitz, Anna Whelan, Nina Ayala, Angela Hardi, Carrie Stoll, Megan Ranney, Cynthia Battle, Michael Silverstein, Emily Miller. The effect of mobile-health interventions on preventing or treating postpartum depression or anxiety: a systematic review and meta-analysis of randomized controlled trials. PROSPERO 2022 CRD42022321649 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022321649 Pim Valentijn, Liza Tymchenko, Maaike Meurs, Matthijs Spruijt, Rosa Y. Arends. The effectiveness of e-mental health for stress, anxiety and depression: A systematic review protocol. PROSPERO 2022 CRD42022311500 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022311500 Zhimin Zheng, Chunxia Li, Rihua Xie, Liqun Yue, Hualing Xie, Jinyu Liao, Yuhua Pan, Xiaoying Chen. Effectiveness of telehealth interventions on post-stroke depression:Systematic Review and Meta-analysis. PROSPERO 2021 CRD42021291311 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021291311 Kuan-Han Lin, Jiun-Yi Wang, Yin-Hwa Shih. Clinical outcomes of telemedicine interventions for patients with depression: a systematic review and meta-analysis. PROSPERO 2021 CRD42021264916 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021264916 Sri Susanty, Made Ari Sarasmita, Yeu-Hui Chuang. Digital intervention to reduce depression during the COVID-19 pandemic: a systematic review. PROSPERO 2021 CRD42021266646 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021266646 Annaleis Giovanetti, Stephen Ilardi, Stephanie Punt. The Acute Efficacy of In-Person Versus Videoconference-Based Psychotherapy for Depression: A Meta-Analysis of Randomized Controlled Trials. PROSPERO 2020 CRD42020156633 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020156633 Buddhika Senanayake, Sumudu Wickramasinghe, Julie Hansen, Mark Chatfield, Anthony Smith, Sisira Edirippulige. The effectiveness of text messaging interventions for depression: a systematic review and meta-analysis. PROSPERO 2019 CRD42019141100 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019141100 Terika McCall, Clinton Bolton III, Rebecca Carlson, Saif Khairat. A systematic review of telehealth interventions for managing anxiety and depression in African American adults. PROSPERO 2018 CRD42018104469 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018104469 Uthara Nair, Sisira Edirippulige, Nigel Armfield, Ruth Crowther. The impact of mHealth, eHealth and telemedicine on maternal depression: a meta-analysis.	11

	PROSPERO 2018 CRD42018093123 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018093123 Christine Rummel-Kluge, Sandra Dietrich, Nicole Koburger. Behavioural and cognitive-behavioural therapy based self-help versus treatment as usual for depression in adults and adolescents [Cochrane Protocol]. PROSPERO 2015 CRD42015027135 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42015027135	

PRIMARY STUDIES

5. Search for primary literature **Date searched: 06-06-22**		
MEDLINE [Ovid MEDLINE(R) ALL 1946 to June 03, 2022]		
#	**Search Statement**	**Results**
1	exp Stress Disorders, Post-Traumatic/ or Depression/ or exp Depressive Disorder/ or exp Anxiety/ or Bipolar Disorder/ or exp Substance-Related Disorders/ or exp Suicide/ or exp Schizophrenia/ or (post?traumatic stress disorder or post?traumatic neuros* or PTSD or moral injury or depression or depressive or melancholia or anxiety or angst or anxieties or hypervigilance or phobia* or panic disorder or obsessive?compulsive disorder or (bipolar adj1 disorder*) or bipolar affective disorder or bipolar affective psychosis or manic?depressive psychosis or manic disorder or substance?related disorder or ((drug or substance or alcohol) adj1 disorder) or ((drug or substance or chemical) adj1 addiction) or alcohol dependen* or suicid* or serious mental illness or SMI or schizophreni*).ti,ab.	1152504
2	exp Telemedicine/ or exp Videoconferencing/ or Electronic Mail/ or exp Telephone/ or Text Messaging/ or exp Mobile Applications/ or exp Cell Phone/ or (telemedicine or telehealth or tele?mental health or tele?behavio?ral health or tele?psychiatry or tele?psychology or tele?therapy or remote or remotely or mHealth or eHealth or (mobile adj1 health) or video?conferenc* or video or VTC or e?mail or online portal or telephone or cell phone or mobile phone or text message or SMS or mobile app*).ti,ab.	350833
3	exp Cognitive Behavioral Therapy/ or Dialectical Behavior Therapy/ or Interpersonal Psychotherapy/ or Implosive Therapy/ or Eye Movement Desensitization Reprocessing/ or Medication Therapy Management/ or Psychiatry/ or (cognitive behavio?ral therap* or cognitive therap* or cognitive behavio?ral conjoint therap* or behavio?r* therap* or exposure therap* or prolonged exposure or narrative exposure or EMDR or eye movement desensitization or motivational interview* or cognitive processing therap* or crisis response plan* or safety plan* or problem?solving therap* or problem?solving skills training or behavio?ral couples therap* or community reinforcement approach or motivational enhancement therap* or twelve?step facilitation or 12?step facilitation or contingency management or brief eclectic psychotherap* or stress inoculation or present?centered therap* or interpersonal therap* or interpersonal psychotherap* or commitment therap* or behavio?ral activation or mindfulness?based cognitive therap* or group therap* or psychodynamic psychotherapy* or couples?focused therap* or social skills training or medication* or pharmacotherapy* or pharmacologic* or psychiatr*).ti,ab.	1023409
4	(teen* or adolescen* or child* or infant* or youth).ti,ab.	2057471
5	1 AND 2 AND 3	5559
6	5 NOT 4	4809
7	limit 6 to yr=”2000-current”, English language	4438
PsycINFO [APA PsycInfo 1806 to May Week 5 2022]		
#	**Search Statement**	**Results**
1	exp Stress Disorders, Post-Traumatic/ or Depression/ or exp Depressive Disorder/ or exp Anxiety/ or Bipolar Disorder/ or exp Substance-Related Disorders/ or exp Suicide/	667305

	or exp Schizophrenia/ or (post?traumatic stress disorder or post?traumatic neuros* or PTSD or moral injury or depression or depressive or melancholia or anxiety or angst or anxieties or hypervigilance or phobia* or panic disorder or obsessive?compulsive disorder or (bipolar adj1 disorder*) or bipolar affective disorder or bipolar affective psychosis or manic?depressive psychosis or manic disorder or substance?related disorder or ((drug or substance or alcohol) adj1 disorder) or ((drug or substance or chemical) adj1 addiction) or alcohol dependen* or suicid* or serious mental illness or SMI or schizophreni*).ti,ab.	
2	exp Telemedicine/ or exp Videoconferencing/ or Electronic Mail/ or exp Telephone/ or Text Messaging/ or exp Mobile Applications/ or exp Cell Phone/ or (telemedicine or telehealth or tele?mental health or tele?behavio?ral health or tele?psychiatry or tele?psychology or tele?therapy or remote or remotely or mHealth or eHealth or (mobile adj1 health) or video?conferenc* or video or VTC or e?mail or online portal or telephone or cell phone or mobile phone or text message or SMS or mobile app*).ti,ab.	112504
3	exp Cognitive Behavioral Therapy/ or Dialectical Behavior Therapy/ or Interpersonal Psychotherapy/ or Implosive Therapy/ or Eye Movement Desensitization Reprocessing/ or Medication Therapy Management/ or Psychiatry/ or (cognitive behavio?ral therap* or cognitive therap* or cognitive behavio?ral conjoint therap* or behavio?r* therap* or exposure therap* or prolonged exposure or narrative exposure or EMDR or eye movement desensitization or motivational interview* or cognitive processing therap* or crisis response plan* or safety plan* or problem?solving therap* or problem?solving skills training or behavio?ral couples therap* or community reinforcement approach or motivational enhancement therap* or twelve?step facilitation or 12?step facilitation or contingency management or brief eclectic psychotherap* or stress inoculation or present?centered therap* or interpersonal therap* or interpersonal psychotherap* or commitment therap* or behavio?ral activation or mindfulness?based cognitive therap* or group therap* or psychodynamic psychotherapy* or couples?focused therap* or social skills training or medication* or pharmacotherapy* or pharmacologic* or psychiatr*).ti,ab.	451320
4	(teen* or adolescen* or child* or infant* or youth).ti,ab.	986797
5	1 AND 2 AND 3	3643
6	5 NOT 4	3095
7	limit 6 to yr=”2000-current”, English language	2712

APPENDIX B: HAND-SEARCHED SYSTEMATIC REVIEWS

1. Bee PE, Bower P, Lovell K, et al. Psychotherapy mediated by remote communication technologies: a meta-analytic review. BMC Psychiatry. 2008;8:60.
2. Bellanti DM, Kelber MS, Workman DE, Beech EH, Belsher BE. Rapid Review on the Effectiveness of Telehealth Interventions for the Treatment of Behavioral Health Disorders. Military Medicine. 2022;187(5-6):e577-e588.
3. Castro A, Gili M, Ricci-Cabello I, et al. Effectiveness and adherence of telephone-administered psychotherapy for depression: A systematic review and meta-analysis. Journal of Affective Disorders. 2020;260:514-526.
4. Coughtrey AE, Pistrang N. The effectiveness of telephone-delivered psychological therapies for depression and anxiety: A systematic review. Journal of Telemedicine & Telecare. 2018;24(2):65-74.
5. Donker T, Petrie K, Proudfoot J, Clarke J, Birch MR, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. Journal of Medical Internet Research. 2013;15(11):e247.
6. Garcia-Lizana F, Munoz-Mayorga I. Telemedicine for depression: a systematic review. Perspectives in Psychiatric Care. 2010;46(2):119-126.
7. Gilmore AK, Wilson SM, Skopp NA, Osenbach JE, Reger G. A systematic review of technology-based interventions for co-occurring substance use and trauma symptoms. Journal of Telemedicine & Telecare. 2017;23(8):701-709.
8. Giovanetti AK, Punt SEW, Nelson EL, Ilardi SS. Teletherapy Versus In-Person Psychotherapy for Depression: A Meta-Analysis of Randomized Controlled Trials. Telemedicine Journal & E Health. 2022;10:10.
9. Guaiana G, Mastrangelo J, Hendrikx S, Barbui C. A Systematic Review of the Use of Telepsychiatry in Depression. Community Mental Health Journal. 2021;57(1):93-100.
10. Harerimana B, Forchuk C, O'Regan T. The use of technology for mental health care delivery among older adults with depressive symptoms: A systematic literature review. International Journal of Mental Health Nursing. 2019;28(3):657-670.
11. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of Synchronous Text-Based Dialogue Systems in Mental Health Interventions: Systematic Review. Journal of Medical Internet Research. 2017;19(8):e267.
12. Hrynyschyn R, Dockweiler C. Effectiveness of Smartphone-Based Cognitive Behavioral Therapy Among Patients With Major Depression: Systematic Review of Health Implications. JMIR MHealth and UHealth. 2021;9(2):e24703.
13. Jiang S, Wu L, Gao X. Beyond face-to-face individual counseling: A systematic review on alternative modes of motivational interviewing in substance abuse treatment and prevention. Addictive Behaviors. 2017;73:216-235.
14. Kiluk BD, Ray LA, Walthers J, Bernstein M, Tonigan JS, Magill M. Technology-Delivered Cognitive-Behavioral Interventions for Alcohol Use: A Meta-Analysis. Alcoholism: Clinical & Experimental Research. 2019;43(11):2285-2295.
15. Kreuze E, Jenkins C, Gregoski M, et al. Technology-enhanced suicide prevention interventions: A systematic review. Journal of Telemedicine & Telecare. 2017;23(6):605-617.
16. Krzyzaniak N, Greenwood H, Scott AM, et al. The effectiveness of telehealth versus face-to face interventions for anxiety disorders: A systematic review and meta-analysis. Journal of Telemedicine & Telecare. 2021:1357633X211053738.

17. Leach LS, Christensen H. A systematic review of telephone-based interventions for mental disorders. Journal of Telemedicine & Telecare. 2006;12(3):122-129.
18. McCall T, Bolton CS, 3rd, Carlson R, Khairat S. A systematic review of telehealth interventions for managing anxiety and depression in African American adults. Began with 2015. 2021;7:31.
19. McCall T, Bolton Iii CS, McCall R, Khairat S. The Use of Culturally-Tailored Telehealth Interventions in Managing Anxiety and Depression in African American Adults: A Systematic Review. Studies in Health Technology & Informatics. 2019;264:1728-1729.
20. McClellan MJ, Osbaldiston R, Wu R, et al. The effectiveness of telepsychology with veterans: A meta-analysis of services delivered by videoconference and phone. Psychological Services. 2022;19(2):294-304.
21. Mohr DC, Vella L, Hart S, Heckman T, Simon G. The Effect of Telephone-Administered Psychotherapy on Symptoms of Depression and Attrition: A Meta-Analysis. Clinical Psychology-Science & Practice. 2008;15(3):243-253.
22. Olthuis JV, Wozney L, Asmundson GJ, Cramm H, Lingley-Pottie P, McGrath PJ. Distance-delivered interventions for PTSD: A systematic review and meta-analysis. Journal of Anxiety Disorders. 2016;44:9-26.
23. Osenbach JE, O'Brien KM, Mishkind M, Smolenski DJ. Synchronous telehealth technologies in psychotherapy for depression: a meta-analysis. Depression & Anxiety. 2013;30(11):1058-1067.
24. Scott AM, Bakhit M, Greenwood H, et al. Real-Time Telehealth Versus Face-to-Face Management for Patients With PTSD in Primary Care: A Systematic Review and Meta-Analysis. Journal of Clinical Psychiatry. 2022;83(4):23.
25. Sloan DM, Gallagher MW, Feinstein BA, Lee DJ, Pruneau GM. Efficacy of telehealth treatments for posttraumatic stress-related symptoms: a meta-analysis. Cognitive Behaviour Therapy. 2011;40(2):111-125.
26. Sullivan SR, Myhre K, Mitchell EL, et al. Suicide and Telehealth Treatments: A PRISMA Scoping Review. Archives of Suicide Research. 2022:1-21.
27. Uhl S, Bloschichak A, Moran A, et al. Telehealth for Substance Use Disorders: A Rapid Review for the 2021 U.S. Department of Veterans Affairs and U.S. Department of Defense Guidelines for Management of Substance Use Disorders. Annals of Internal Medicine. 2022;175(5):691-700.

APPENDIX C: EXCLUDED STUDIES

Exclude reasons: 1=Ineligible population, 2=Ineligible intervention, 3=Ineligible comparator, 7=Ineligible publication type, 8=Outdated or ineligible systematic review.

Citation	Exclude Reason
Andreasson K, Krogh J, Bech P, et al. MYPLAN -mobile phone application to manage crisis of persons at risk of suicide: study protocol for a randomized controlled trial. *Trials [Electronic Resource].* 2017;18(1):171.	E2
Axelsson E, Andersson E, Ljotsson B, Bjorkander D, Hedman-Lagerlof M, Hedman-Lagerlof E. Effect of internet vs face-to-face cognitive behavior therapy for health anxiety: A randomized noninferiority clinical trial. *JAMA Psychiatry.* 2020;77(9):915-924.	E2
Boykin DM, Keegan F, Thompson KE, Voelkel E, Lindsay JA, Fletcher TL. Video to Home Delivery of Evidence-Based Psychotherapy to Veterans With Posttraumatic Stress Disorder. *Frontiers in psychiatry Frontiers Research Foundation.* 2019;10:893.	E3
Callan JA, Howland RH, Puskar K. Using computers and the Internet for psychiatric nursing intervention. *Journal of Psychosocial Nursing & Mental Health Services.* 2009;47(1):13-14.	E7
Chae YM, Park HJ, Cho JG, Hong GD, Cheon KA. The reliability and acceptability of telemedicine for patients with schizophrenia in Korea. *Journal of Telemedicine & Telecare.* 2000;6(2):83-90.	E2
DeFulio A, Rzeszutek MJ, Furgeson J, Ryan S, Rezania S. A smartphone-smartcard platform for contingency management in an inner-city substance use disorder outpatient program. *Journal of Substance Abuse Treatment.* 2021;120:108188.	E2
Drugs CAf, Health Ti. CADTH optimal use reports. In: *Use of Solvent/Detergent-Treated Human Plasma (Octaplas): Pilot Project.* Canadian Agency for Drugs and Technologies in Health, Ottawa (ON); 2011.	E8
Egede LE, Frueh CB, Richardson LK, et al. Rationale and design: telepsychology service delivery for depressed elderly veterans. *Trials [Electronic Resource].* 2009;10:22.	E7
Ekberg J, Timpka T, Bang M, Froberg A, Halje K, Eriksson H. Cell phone-supported cognitive behavioural therapy for anxiety disorders: a protocol for effectiveness studies in frontline settings. *BMC Medical Research Methodology.* 2011;11:3.	E7
Fortier CB, Currao A, Kenna A, et al. Online Telehealth Delivery of Group Mental Health Treatment Is Safe, Feasible, and Increases Enrollment and Attendance in Post-9/11 U.S. Veterans. *Behavior Therapy.* 2022;53(3):469-480.	E1
Fortney JC, Pyne JM, Mouden SB, et al. Practice-based versus telemedicine-based collaborative care for depression in rural federally qualified health centers: a pragmatic randomized comparative effectiveness trial. *American Journal of Psychiatry.* 2013;170(4):414-425.	E2
Fortney JC, Pyne JM, Mouden SB, et al. Practice-Based Versus Telemedicine-Based Collaborative Care for Depression in Rural Federally Qualified Health Centers: A Pragmatic Randomized Comparative Effectiveness Trial. *Focus.* 2017;15(3):361-372.	Duplicate
Gilmore AK, Davis MT, Grubaugh A, et al. "Do you expect me to receive PTSD care in a setting where most of the other patients remind me of the perpetrator?": Home-based telemedicine to address barriers to care unique to military sexual trauma and veterans affairs hospitals. *Contemporary Clinical Trials.* 2016;48:59-64.	E7

Citation	Exclude Reason
Glueckauf RL, Davis WS, Willis F, et al. Telephone-based, cognitive-behavioral therapy for African American dementia caregivers with depression: initial findings. *Rehabilitation Psychology.* 2012;57(2):124-139.	E1
Gratzer D, Khalid-Khan F. Internet-delivered cognitive behavioural therapy in the treatment of psychiatric illness. *CMAJ Canadian Medical Association Journal.* 2016;188(4):263-272.	E8
Gratzer D, Khalid-Khan F, Balasingham S, Yuen N, Jayanthikumar J. Internet-delivered cognitive behavioural therapy in a Canadian community hospital: A novel approach to an evidence-based intervention. *Canadian Journal of Community Mental Health.* 2018;37(1):81-85.	E2
Gros DF, Strachan M, Ruggiero KJ, et al. Innovative service delivery for secondary prevention of PTSD in at-risk OIF-OEF service men and women. *Contemporary Clinical Trials.* 2011;32(1):122-128.	E7
Harder VS, Musau AM, Musyimi CW, Ndetei DM, Mutiso VN. A randomized clinical trial of mobile phone motivational interviewing for alcohol use problems in Kenya. *Addiction.* 2020;115(6):1050-1060.	E2
Johansson M. Treating alcohol use disorder on the internet. *Dissertation Abstracts International: Section B: The Sciences and Engineering.* 2021;82(9-B).	E2
Kay-Lambkin FJ, Baker AL, Palazzi K, Lewin TJ, Kelly BJ. Therapeutic Alliance, Client Need for Approval, and Perfectionism as Differential Moderators of Response to eHealth and Traditionally Delivered Treatments for Comorbid Depression and Substance Use Problems. *International Journal of Behavioral Medicine.* 2017;24(5):728-739.	E2
Kiropoulos LA, Klein B, Austin DW, et al. Is internet-based CBT for panic disorder and agoraphobia as effective as face-to-face CBT? *Journal of Anxiety Disorders.* 2008;22(8):1273-1284.	E2
Kishimoto T, Kinoshita S, Bun S, et al. Japanese Project for Telepsychiatry Evaluation during COVID-19: Treatment Comparison Trial (J-PROTECT): Rationale, design, and methodology. *Contemporary Clinical Trials.* 2021;111:106596.	E7
Leon-Salas A, Hunt JJ, Richter KP, Nazir N, Ellerbeck EF, Shireman TI. Pharmaceutical assistance programs to support smoking cessation medication access. *Journal of the American Pharmacists Association: JAPhA.* 2017;57(1):67-71.e61.	E2
Leterme AC, Behal H, Demarty AL, et al. A blended cognitive behavioral intervention for patients with adjustment disorder with anxiety: A randomized controlled trial. *Internet Interventions.* 2020;21:100329.	E2
Lundstrom L, Flygare O, Andersson E, et al. Effect of Internet-Based vs Face-to-Face Cognitive Behavioral Therapy for Adults With Obsessive-Compulsive Disorder: A Randomized Clinical Trial. *JAMA Network Open.* 2022;5(3):e221967.	E2
Luxton DD, Pruitt LD, O'Brien 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. *Contemporary Clinical Trials.* 2014;38(1):134-144.	E7
Ly KH, Topooco N, Cederlund H, et al. Smartphone-Supported versus Full Behavioural Activation for Depression: A Randomised Controlled Trial. *PLoS ONE [Electronic Resource].* 2015;10(5):e0126559.	E2
Mathiasen K, Andersen TE, Riper H, Kleiboer AA, Roessler KK. Blended CBT versus face-to-face CBT: a randomised non-inferiority trial. *BMC Psychiatry.* 2016;16(1):432.	E7

Citation	Exclude Reason
McKay JR. Is there a case for extended interventions for alcohol and drug use disorders? *Addiction.* 2005;100(11):1594-1610.	E3
Morland LA, Pierce K, Wong MY. Telemedicine and coping skills groups for Pacific Island veterans with post-traumatic stress disorder: A pilot study. *Journal of Telemedicine and Telecare.* 2004;10(5):286-289.	E2
Nicholas J, Knapp AA, Vergara JL, et al. An Exploratory Brief Head-To-Head Non-Inferiority Comparison of an Internet-Based and a Telephone-Delivered CBT Intervention for Adults with Depression. *Journal of Affective Disorders.* 2021;281:673-677.	E2
Peterson AL, Resick PA, Mintz J, et al. Design of a clinical effectiveness trial of in-home cognitive processing therapy for combat-related PTSD. *Contemporary Clinical Trials.* 2018;73:27-35.	E7
Romijn G, Batelaan N, Koning J, et al. Acceptability, effectiveness and cost-effectiveness of blended cognitive-behavioural therapy (bCBT) versus face-to-face CBT (ftfCBT) for anxiety disorders in specialised mental health care: A 15-week randomised controlled trial with 1-year follow-up. *PloS one.* 2021;16(11):e0259493.	E2
Santesteban-Echarri O, Piskulic D, Nyman RK, Addington J. Telehealth interventions for schizophrenia-spectrum disorders and clinical high-risk for psychosis individuals: A scoping review. *Journal of Telemedicine & Telecare.* 2020;26(1-2):14-20.	E8
Sarkar S, Gupta R. Telephone vs face-to-face cognitive behavioral therapy for depression. *JAMA.* 2012;308(11):1090-1091; author reply 1091.	E7
Smucker Barnwell SV, Juretic MA, Hoerster KD, Van de Plasch R, Felker BL. VA Puget Sound Telemental Health Service to rural veterans: a growing program. *Psychological Services.* 2012;9(2):209-211.	E3
Strachan M, Gros DF, Yuen E, Ruggiero KJ, Foa EB, Acierno R. Home-based telehealth to deliver evidence-based psychotherapy in veterans with PTSD. *Contemporary Clinical Trials.* 2012;33(2):402-409.	E7
Thase ME, McCrone P, Barrett MS, et al. Improving Cost-effectiveness and Access to Cognitive Behavior Therapy for Depression: Providing Remote-Ready, Computer-Assisted Psychotherapy in Times of Crisis and Beyond. *Psychotherapy & Psychosomatics.* 2020;89(5):307-313.	E2
Wagner B, Horn AB, Maercker A. Internet-based versus face-to-face cognitive-behavioral intervention for depression: A randomized controlled non-inferiority trial. *Journal of Affective Disorders.* 2014;152(154):113-121.	E2
Watts S, Mackenzie A, Thomas C, et al. CBT for depression: a pilot RCT comparing mobile phone vs. computer. *BMC Psychiatry.* 2013;13:49.	E2
Wei N, Huang BC, Lu SJ, et al. Efficacy of internet-based integrated intervention on depression and anxiety symptoms in patients with COVID-19. *Journal of Zhejiang University SCIENCE B.* 2020;21(5):400-404.	E1
Wright JH, Wright AS, Albano AM, et al. Computer-Assisted Cognitive Therapy for Depression: Maintaining Efficacy While Reducing Therapist Time. *The American Journal of Psychiatry.* 2005;162(6):1158-1164.	E2

APPENDIX D: EVIDENCE TABLES

CHARACTERISTICS OF INCLUDED PRIMARY STUDIES

RCTs

Author Year N Country	Follow-up	Participant Characteristics	Intervention	Telehealth Modality Telehealth Setting	Comparator Modality Comparator Setting
Posttraumatic stress disorder					
Acierno 2016[1] & Strachan 2012[2] N=265 US	Post-treatment, 3 months, 12 months	Veterans with PTSD Mean age: 45.6 years 94.4% male 50.4% White, 47.4% Black	BA-TE	Videoconference Home	In-person Clinic
Acierno 2017[3] & Gros 2018[4] & Gros 2018[5] & Reich 2021[6] & Yuen 2015[7] N=150 US	Post-treatment, 3 months, 6 months	Veterans with PTSD Mean age: 41.8 years 96.2% male 60.6% White, 33.3% Black	PE	Videoconference Home	In-person Clinic

Author Year N Country	Follow-up	Participant Characteristics	Intervention	Telehealth Modality Telehealth Setting	Comparator Modality Comparator Setting
Acierno 2021[8] & Gilmore 2020[9] & Lopez 2021[10] N=136 US	Post-treatment, 3 months, 6 months	Women Veterans with MST-related PTSD Mean age: 43.4 years 0% male 63.97% Black, 29.41% White	PE	Videoconference Home	In-person Clinic
Franklin 2017[11] N=27 US	Post-treatment, 1 month	Veterans with PTSD Mean age: 46.1 years 92.6% male 69.2% White, 23.1% Black	PE	Videoconference (iPhone) Home	Videoconference Clinic
Frueh 2007[12] & Frueh 2007[13] N=38 US	Post-treatment, 3 months	Veterans with PTSD Mean age: 55.2 years 100% male 42.1% white	CBT	Videoconference Clinic	In-person Clinic
Glassman 2019a[14] N=251 US	Post-treatment, 3 months, 6 months	Veterans with PTSD Mean age: 50.87 years 49.8% male Race NR	Individual or group CPT	Videoconference Clinic	In-person Clinic

Author Year N Country	Follow-up	Participant Characteristics	Intervention	Telehealth Modality Telehealth Setting	Comparator Modality Comparator Setting
Haghnia 2019[15] N=71 Iran	Post-treatment	Veterans with PTSD Mean age: 50 years 100% male Race NR	Psychiatry	Any/multiple Home	In-person Clinic
Hernandez-Tejada 2014$_b$[16] N=258 US	NR	Veterans with PTSD Mean age: 46.5 years 100% male 57.8% White	PE	Videoconference Home	In-person Clinic
Liu 2020[17] N=207 US	Post-treatment, 6 months	Veterans with PTSD Mean age: 48.4 years 77.4% male 55.2% Caucasian, 20.6% Hispanic, 15.5% Black	CPT	Videoconference Clinic	In-person Clinic
Maieritsch, 2016[18] N=90 US	Post-treatment, 3 months	Veterans with PTSD Mean age: 30.9 93.3% male Race NR	CPT	Videoconference Clinic	In-person Clinic

Author Year N Country	Follow-up	Participant Characteristics	Intervention	Telehealth Modality Telehealth Setting	Comparator Modality Comparator Setting
Morland 2014[19] & Morland 2011[20] N=125 US	Post-treatment, 3 months, 6 months	Veterans with PTSD Mean age: 55.3 years 100% male 46.4% White	Group CPT	Videoconference Clinic	In-person Clinic
Morland 2015[21] N=126 US	Post-treatment, 3 months, 6 months	Recruited at Veteran sites (but mostly civilians) with PTSD Mean age: 46.4 years 0% male 47.6% White % Black NR	CPT	Videoconference Clinic	In-person Clinic
Morland 2020[22] N=175 US	Post-treatment, 3 months, 6 months	Veterans with PTSD Mean age: 46.5 years 74.9% male 40.6% White 28.6% Black	PE	Videoconference Home, clinic	In-person Home
Morland 2022[23] N=137 US	Post-treatment, 3 months, 6 months	Veteran couples with PTSD Mean age: 40.9 years 80.4% male (Veterans) 55.5% White 20.5% Black	Brief cognitive-behavioral conjoint therapy	Videoconference Home	In-person Clinic

Author Year N Country	Follow-up	Participant Characteristics	Intervention	Telehealth Modality Telehealth Setting	Comparator Modality Comparator Setting
Peterson 2022[24] N=120 US	Post-treatment, 6 months	Active-duty military or Veterans with PTSD Mean age: 40.5 years 88% male 37% White 17% Black	CPT	Videoconference Home	In-person Clinic, home
White 2021b[25] N=140 US	Post-treatment, 3 months, 6 months	Veterans with PTSD Mean age: 42.4 years 63.6% female 40.7% White 51.4% Black	PE	Videoconference Home	In-person Clinic
Ziemba 2014[26] N=18 US	Post-treatment	Veterans with PTSD Age NR 90% male 79% Black	CBT	Videoconference Clinic	In-person Clinic

Author Year **N** **Country**	**Follow-up**	**Participant Characteristics**	**Intervention**	**Telehealth Modality** **Telehealth Setting**	**Comparator Modality** **Comparator Setting**
Depression					
Alegria 2014[27] & Aguilera 2018[28] & Alcantara 2016[29] & Kafali 2014[30] N=257 US	2 months, 4 months	Low-income Latinos with depression 26.3% 18-34, 36.3% 35-49, 34.5% 50-64, 2.9% ≥65 81.3% female 33.3% Black or dark-skinned, 26.9% White	Engagement and counseling for Latinos (CBT plus care management)	Telephone Home	In-person Clinic
Choi 2014[31] N=121 US	12 weeks, 24 weeks, 36 weeks	Low-income homebound older adults with depression symptoms Mean age: 65.21 years 22.3% male 41.3% white 33.9% Black	Problem-solving therapy – primary care version	Videoconference Home	In-person Home
Egede 2015[32] & Egede 2016[33] & Egede 2017[34] & Egede 2018[35] N=241 US	Mid-treatment, post-treatment, 3 months, 12 months	Older Veterans with MDD Mean age: 63.9 years 97.5% male 59.3% White, 39.7% Black	Behavioral activation for depression	Videoconference Home	In-person Clinic

Author Year **N** **Country**	**Follow-up**	**Participant Characteristics**	**Intervention**	**Telehealth Modality** **Telehealth Setting**	**Comparator Modality** **Comparator Setting**
Hungerbuehler 2016[36] N=107 Brazil	6 months, 12 months	Adults with mild depression Mean age: 35.64 years 29% male 96.3% Brazilian	Psychiatry	Videoconference Home	In-person Clinic
Luxton 2016[37] & Bounthavong 2018[38] & Pruitt 2019[39] & Smolenski 2017[40] N=121 US	Mid-treatment, post-treatment, 3 months	Military service members and Veterans with depression Mean age: 35.2 years 81.8% male 70.2% White	Behavioral activation	Videoconference Home	In-person Clinic
Mohr 2012[41] & Kalapatapu 2014[42] & Stiles-Shields 2014[43] & Stiles-Shields 2015[44] N=325 US	Post-treatment, 3 months, 6 months	Primary care patients with depression Mean age: 47.7 years 22.5% male 57.5% White	CBT	Telephone Home	In-person Clinic

Author Year **N** **Country**	**Follow-up**	**Participant Characteristics**	**Intervention**	**Telehealth Modality** **Telehealth Setting**	**Comparator Modality** **Comparator Setting**
Ruskin 2004[45] & Rhein 2001[46] N=131 US	Post-treatment	Veterans with depression Mean age: 49.7 years 85.6% male (reported only for subset of n=90 participants)	Psychiatry	Videoconference Clinic	In-person Clinic
Anxiety					
Cherestal 2019[47] N=22 US	Post-treatment only	Adults with specific phobia (aviophobia) For 18 participants included in analyses: Mean age: 46.78 years 33% male 88.89% White	Virtual reality exposure therapy	Telephone Home	In-person Clinic
Lovell 2006[48] N=72 UK	Post-treatment, 3 months, 6 months	Adults with OCD diagnosis Mean age: 31.9 years 40.3% male Race NR	Exposure and response prevention (ERP)	Telephone Home	In-person Clinic
Watts 2020[49] N=115 Canada		Adults with anxiety Mean age: 41.5 years 17.4% male 100% White	CBT	Videoconference Clinic	In-person Clinic

Author Year N Country	Follow-up	Participant Characteristics	Intervention	Telehealth Modality Telehealth Setting	Comparator Modality Comparator Setting
Multiple mental health conditions					
De Las Cuevas 2006[50] N=140 Spain	Post-treatment only	Adults with ICD-10 diagnosis of F1, F2, F3, F4, and F6 17.86% <25, 50% 25-45, 26.43% 45-65, 5.71% > 65 33.57% male Race NR	Psychiatry	Videoconference Clinic	In-person Clinic
O'Reilly 2007[51] N=495 Canada	Post-treatment, 4 months, 1 year	Adults with any condition eligible for referral to psychiatric clinic Median age between 35 and 44 years 37% male Race NR	Psychiatry	Videoconference Home	In-person Clinic
Stubbings 2013[52] N=26 Australia	Post-treatment, 1.5 months	Adults living in Perth, Australia with mood or anxiety disorder Mean age: 30 years 42.3% male Race NR	CBT	Videoconference Clinic	In-person Clinic

Notes. [a] *Secondary analysis of data from 2 RCTs (Morland 2014; Morland 2015)* [b] *Secondary analysis of data from 2 RCTs (Acierno 2017; Acierno 2021). Abbreviations. BA-TE=Behavioral Activation and Therapeutic Exposure; CBT=cognitive behavioral therapy; CPT=cognitive processing therapy; ICD=International Statistical Classification of Diseases; MDD=major depressive disorder; MST=military sexual trauma; NR=not reported; OCD=obsessive-compulsive disorder; PE=prolonged exposure; PTSD=posttraumatic stress disorder; RCT=randomized control trial.*

Cohort Studies

Author Year N Country	Follow-up	Participant Characteristics	Intervention	Telehealth Modality Telehealth Setting	Comparator Modality Comparator Setting
Posttraumatic stress disorder					
Germain 2009[53] & Germain 2010[54] N=48 Canada	Post-treatment	Adults with PTSD Mean age: 42.33 years 39.58% male Race NR	CBT	Videoconference Clinic	In-person Clinic
Gros 2011[55] N=89 US	Post-treatment	Veterans with PTSD For telehealth group: Mean age: 45.1 93.5% male 50% White, 45.2% Black	Exposure therapy	Videoconference Clinic	In-person Clinic
Knowlton 2021[56] N=581 US	Post-treatment	Veterans with PTSD Mean age: 47.14 years 79.2% male 83.5% European American	PE, CPT	Videoconference Clinic or home	In-person Clinic
LoSavio 2021[57] N=277 US	Post-treatment	Veterans with PTSD Mean age: 47.79 years 76.2% male 52.4% White, 34.6% Black	Written exposure therapy (WET)	Videoconference NR	In-person Clinic

Author **Year** **N** **Country**	**Follow-up**	**Participant Characteristics**	**Intervention**	**Telehealth Modality** **Telehealth Setting**	**Comparator Modality** **Comparator Setting**
Tuerk 2010[58]	Post-treatment	Veterans with PTSD	PE	Videoconference	In-person
N=47		Mean age: 39 years 94% male		Clinic	Clinic
US		64% White 34% Black			
Valentine 2020[59]	Post-treatment	Veterans with MST-related PTSD	CPT or PE	Videoconference	In-person
N=171		Mean age: 44.4 years old 26.5% male		Home	Clinic
US		68.5% White 22.8% Black			
Wierwille 2016[60]	Post-treatment	Veterans with PTSD	CPT or PE	Videoconference	In-person
N=221		Mean age: 46.74 years 87.1% male		Clinic	Clinic
US		85.5% White % Black NR			
Depression					
Ritchie 2007[61]	Post-treatment	Retired or active-duty military service members with depression (and their families)	Group CBT	Videoconference	In-person
N=14				Clinic	Clinic
US		Mean age: 33.50 years 35.7% male Race NR			

Author Year **N** **Country**	**Follow-up**	**Participant Characteristics**	**Intervention**	**Telehealth Modality** **Telehealth Setting**	**Comparator Modality** **Comparator Setting**
Anxiety					
Bouchard 2004[62] N=21 Canada	Post-treatment, 6 months	Adults diagnosed with panic disorder with agoraphobia Mean age: 37.99 years 28.57% male Race NR	CBT	Videoconference Clinic	In-person Clinic
Bouchard 2020[63] N=71 Canada	Post-treatment, 12 months	Adults diagnosed with panic disorder with agoraphobia Mean age: 35.77 years 17% male Race NR	CBT	Videoconference Clinic	In-person Clinic
Milosevic 2022[64] N=413 Canada	Post-treatment, 3 months	Adults with anxiety or panic disorder/agoraphobia diagnosis Mean age: 34.9 years 31.6% male 87.4% White % Black NR	Group CBT	Videoconference Home	In-person Clinic
Pinciotti 2022[65] N=468 US	Post-treatment	Adults with OCD Mean age: 29.9 years 47.4% male 73.5% White % Black NR	CBT/exposure and response prevention	Videoconference Home	In-person Clinic

Author **Year** **N** **Country**	**Follow-up**	**Participant Characteristics**	**Intervention**	**Telehealth Modality** **Telehealth Setting**	**Comparator Modality** **Comparator Setting**
Substance use					
Bean 2022[66] N=69 US	Post-treatment	Adults with co-occurring SUD and at least one mental health disorder (most commonly depression, anxiety, or PTSD) Mean age: 40.51 years 73.91% male 94.2% White, 2.9% Black	DBT-based IOP	Videoconference Home	In-person Clinic
Vakkalanka 2022[67] N=28,791 US	Post-treatment	Veterans with OUD 48.8% between 25-44 years old 92.9% male 81.2% White 12.5% Black	Counseling to accompany buprenorphine prescription	Videoconference Any/multiple	In-person Clinic
Zheng 2017[68] N=100 US	Variable	People with OUD Mean age: 35.7 years 46% male 94% White 3% Black	Medication-assisted treatment	Videoconference Clinic	In-person Clinic

Author Year N Country	Follow-up	Participant Characteristics	Intervention	Telehealth Modality Telehealth Setting	Comparator Modality Comparator Setting
Multiple mental health conditions					
Gannon 2021[69] N=391 US	Post-treatment	Adults with depressive disorders, bipolar disorders, anxiety disorders, borderline personality disorder, PTSD, hallucinogen use, and other psychotic, eating, and adjustment disorders Mean age: 33.1 years 36% male 76% White, 17% Black	Transdiagnostic psychiatric IOP	Videoconference Home	In-person Clinic
Hammond 2012[70] N=4,106 UK	Post-treatment	Adults in mental health treatment 16% <18-25, 24% 26-35, 25% 36-45, 19% 46-55, 12% 56-65, 4% 66-85 35% male Race NR	Low-intensity CBT	Telephone Home	In-person Clinic
Khatri 2014[71] N=18 Canada	3-month	Adults with mood, anxiety, or adjustment disorder diagnosis Mean age: 50.8 years 28% male 100% White	Group CBT	Videoconference Home	In-person Clinic
Liou 2022[72] N=288 US	Post-treatment	Adults with a psychiatric diagnosis Mean age: 51.4 years 29.5% male Race NR	EMDR + CBT	Videoconference Home	In-person Clinic

Author Year **N** **Country**	**Follow-up**	**Participant Characteristics**	**Intervention**	**Telehealth Modality** **Telehealth Setting**	**Comparator Modality** **Comparator Setting**
Zimmerman 2021[73] N=414 US	Post-treatment	Adults with a variety of presenting mental health concerns Mean age: 37.0 years 28.3% male 72.9% White 7.5% Black	Partial hospitalization program based on ACT and related evidence-based psychotherapy techniques	Videoconference Home	In-person Clinic

Abbreviations. ACT=acceptance and commitment therapy; CBT=cognitive behavioral therapy; CPT=cognitive processing therapy; DBT=dialectical behavior therapy; EMDR=eye movement desensitization and reprocessing; IOP=intensive outpatient program; MDD=major depressive disorder; MST=military sexual trauma; NR=not reported; OCD=obsessive-compulsive disorder; OUD=opioid use disorder; PE=prolonged exposure; PTSD=post-traumatic stress disorder; SUD=substance use disorder.

OUTCOME DATA OF INCLUDED PRIMARY STUDIES

RCTs

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
PTSD					
Acierno 2016[1] & Strachan 2012[2]	PCL-M, BDI-II, BAI The lower bound of the CI for the between treatment difference in mean PCL scores for VTC-H relative to IP-C scores were well within the prespecified range of the meaningful clinical difference (−Δ = 8.8), with Δ = −0.11 at PT, Δ = −1.84 at month 3, and Δ = −0.66 at month 12. For mean BDI scores, outcomes of VTC-H treatment effect sizes were well within the prespecified range (−Δ = 5), with Δ = 0.89 at PT, Δ = 1.18 at 3-month follow-up, and Δ = −0.29 at 12-month follow-up. BAI results are only presented for a subset of 31 participants at PT and no statistically significant differences were found between groups (F = .3, *p* = .594, *d* = .23)	Did not complete at least 5 sessions and provide any posttreatment or follow-up data. 23.1% for IP-C, 18.0% for VTC-H (X^2 = 0.93, *p* = .212)	None N/A	None N/A	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Acierno 2017[3] & Gros 2018[4] & Gros 2018[5] & Reich 2021[6] & Yuen 2015[7]	PCL-M, CAPS, BDI-II, BAI, IPF The lower limits of the 90% CI for PCL scores at PT (*M* = −3.2; 90% CI: −8.6 to 2.1), 3 months (*M* = −2.8; 90% CI: −7.6 to 2.0), and 6 months (*M* = 0.03; 90% CI: −4.9 to 5.0) were within the pre-specified meaningful clinical difference (−Δ = −8.8), indicating that VTC-H was "as good as" IP-C. The lower limit of the 90% CI for BDI scores at PT (*M* = −2.4; 90% CI: −6.3 to 1.5) and 3-month follow-up (*M* = −2.0; 90% CI: −5.7 to 1.6) fell slightly outside the pre-specified limit of non-inferiority (Δ = −5.0). However, scores between conditions were virtually identical (*M* = −0.3; 90% CI: −4.1 to 3.6) at 6-month follow-up. There were no significant differences on IPF pre-treatment, PT, and change score between modality (*t* = 0.764, *p* = 0.445). CAPS and BAI follow-up data are only provided in preliminary findings (n = 52). There were no significant differences among the rates of diagnosis across the conditions (X^2 = 0.62, *p* = 0.73). At PT, for the CAPS, the	Did not complete at least 6 therapy sessions. 19% of IP-C group, 32.8% of VTC-H group (X^2 = 3.23, *P* = 0.072). There were no statistically significant differences in the average number of sessions completed prior to dropout between conditions in the dropout sample (F(1,131) = 0.07, *p* = .80).	Session attendance. predictors of dropout There were no statistically significant differences in the overall average number of sessions received in both groups (F(1,131) = 2.97, *p* = .09) in the overall ITT sample. In discrete time survival analysis, treatment condition was examined as a baseline predictor of treatment discontinuation and significantly predicted treatment discontinuation with an odds ratio of 1.97 (95% CI: 1.02, 3.82). The discontinuation rate was greater in the VTC-H condition compared with the IP-C condition.	SDPQ, CPOSS There was a significant effect of treatment modality on the SDPQ telehealth travel item (F = 5.1; *p* = 0.029), with participants in the VTC-H group endorsing willingness to travel slightly further for telehealth services than participants in the IP-C condition. There were no significant effects of modality on any of the measures of the perception of the quality of service delivery and satisfaction with services provided (F > 2.1; *p* < 0.16; partial η^2 < 0.040).	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
	ITT effect size (*g* = 0.13; 90% CI [0.32, 0.59]) and per protocol effect size (g = 0.12; 90% CI [−0.36, 0.60]) supported noninferiority. For the BAI, the ITT effect size (g = 0.10; 90% CI [−0.36, 0.55]) and per protocol effect size (*g* = 0.12; 90% CI [−0.42, 0.59]) supported noninferiority.				

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Acierno 2021[8] & Gilmore 2020[9] & Lopez 2021[10]	PCL-5, BDI-II There were no differences between IP-C and VTC-H delivery on the PCL-5 at PT (ITT: b = 1.91, *p* = .381; completer: b = 4.09, *p* = 1.95), nor did the rate of change in PTSD symptoms during (ITT: b = 0.10, *p* = .673; completer: *b* = 0.06, *p* = .826) and after treatment differ by mode of delivery (ITT: *b* = 0.05, *p* = .639; completer: *b* = 0.03, *p* = .845). Within the ITT sample, VTC-H delivery was associated with significantly more depressive symptoms than IP-C delivery at the PT assessment (ITT: b = 4.03 *p* = .044). However, amongst treatment completers, VTC-H did not differ from IP-C delivery in PT depression symptoms (completer: *b* = 4.50, *p* = .099). The rate of change in depressive symptoms during (ITT: *b* = 0.13, *p* = .411; completer: *b* = 0.03, *p* = .879) and after treatment did not differ by the mode of delivery (ITT: *b* = 0.01, *p* = .848; completer: *b* = - 0.04, *p* = .606). Treatment condition was not a significant predictor of PTSD symptoms (*B*	Did not complete 8 or more sessions 49.3% of participants in each condition completed treatment. Treatment condition was not associated with dropout (*B* = 0.199, *SE* = 0.383, OR = 1.22, *p* = .603).	Session attendance, predictors of dropout Participants assigned to IP-C delivery of PE completed an average of 6.28 (SD = 4.33) sessions, whereas those assigned to VTC-H delivery of PE completed an average of 6.80 (SD = 4.14) sessions. Dose received did not differ between VTC-H and IP-C delivery of PE, t(134) = - 0.71, *p* = .481. The only significant predictor of treatment dropout was difficulties with emotion regulation (odds ratio, 1.03; *p* < .01). Posttraumatic stress symptoms and depression diagnosis were not significant predictors.	None N/A	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
	= .03, *SE* = .16, *t* = .22, *p* = .826) or depression symptoms (*B* = .06, *SE* = .13, *t* = .50, *p* = .618).				
Franklin 2017[11]	CAPS, PDS, BDI-II, BAI No significant differences were found between groups on the PDS, BDI, BAI, or CAPS. At 1-month follow-up, all participants in the VTC-H group had their PTSD symptoms offset below DSM-IV-TR PTSD diagnostic cutoffs, whereas one (33.3%) participant in VTC-C had their PTSD offset on the CAPS.	Did not complete treatment 60% completed treatment. Differences in dropout rates between groups were statistically significant, $X^2(2) = 8.333$, $p = .016$.	Documentation of issues with teleconferencing equipment 20% (2) of the 10 Veterans in the VTC-H group had poor cell service in their home area to the extent that it was interfering with therapy, and both dropped out of treatment. Another 20% (2) Veterans did not have quiet, undisturbed rooms in which to participate in the therapy.	Acceptability (asked patients which treatment they prefer prior to randomization) Most of the 27 randomized veterans preferred to receive PE via VTC-H (n = 14, 51.9%) rather than in-person, (n = 6, 22.2%), VTC-C (n = 3, 11.1%), or TAU (n = 1, 3.7%); others had no preference (n = 3, 11.1%). Treatment preference did not differ significantly across the assigned treatment groups, $X^2(8) = 7.40$, $p = .49$.	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Frueh 2007[12] & Frueh 2007[13]	PCL-M, BDI, SCL-90-R GSI Noninferiority analysis indicated that VTC-C was not inferior to IP-C at post-treatment or 3-month follow-up. Using a standard superiority hypothesis testing approach, no significant group differences were found for change scores on clinical outcome measures from pre-treatment to post-treatment or 3-month follow-up. The maximum amount by which VTC-C was inferior to IP-C improved somewhat for change in PTSD severity and for social activities outside the home.	Ceased treatment before end of scheduled 14 weeks For the VTC-C group, 8 participants (47%) did not complete the active treatment phase compared to 9 participants (43%) in the IP-C ($p > 0.05$).	Session attendance, homework completion No significant group differences were found for session attendance (IP-C 8.48 vs VTC-C 6.41, $p = 0.21$). However, the IP-C group was more likely to have completed homework assignments ($p = 0.04$).	CPOSS-VA, TCS No statistically significant differences in CPOSS-VA or TCS results.	None N/A
Glassman 2019[14]	QOLI In the model investigating the impact of treatment modality on QOLI scores with the growth factors equated across studies, treatment modality did not predict any of the growth factors, with effect estimates for the intercept estimate = -0.27 (-0.81, 0.27), slope estimate = 0.004 (-0.60, 0.61) and quadratic estimate =0.001 (-0.18, 0.20); all *p*s > 0.33.	Did not complete study protocol Approximately 16% of women and 14% of men dropped out of treatment prior to completing the study protocol.	None N/A	None N/A	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Haghnia 2019[15]	None N/A	Did not complete intervention 5 of the patients did not receive intervention and 3 more discontinued the intervention (1 in IP-C and 2 in VTC-H).	Session attendance, access to a psychiatrist & wait times (study questionnaire) Compared with the IP-C group, the VTC-H group had significantly greater completion of therapy sessions (t = 7.19, df = 59, p = 0.001). There were no significant differences between the groups on access to psychiatrists (t = 1.22, df = 59, p = 0.227) or patients' wait times (t = -1.2, df = 59, p = 0.231).	Patient satisfaction, treatment cost (study questionnaire) Compared with IP-C, VTC-H elevated patient satisfaction (t = 3.85, df = 59, p = 0.002 and had lower treatment costs (t = 5.996, df = 59, p = 0.001).	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Hernandez-Tejada 2014[16]	None N/A	Did not complete treatment Rate of dropout from VTC-H (28.7%, n = 35) was not significantly different from that of IP-C treatment (25.0%, n = 34; X^2 = 0.45, p = 0.30).	Barriers to Exposure Therapy Participation Scale (BTPS) For the BTPS, considering the Stressors and Obstacles factor, 1-way ANOVA indicated a significant difference between IP-C and VTC-H participants, with IP-C participants reporting more problems with bad weather, parking, transportation, and work/family obligations ($\bar{x}$ = 22.5) compared to VTC-H participants ($\bar{x}$ = 19.6; F(1, 45) = 5.20, p = 0.027). No other significant differences were noted for any other factor.	TAQ Significant differences were found between groups with respect to reporting that they would feel comfortable using telemedicine at a local church, with a greater proportion of IP-C (72.2%) vs VTC-H participants (41.7%) reporting comfort. Similarly, 55.6% of IP-C participants felt comfortable receiving telemedicine services at a local clinic, compared to only 25% of those who were in the VTC-H condition. Majorities of both groups indicated that telemedicine would not be as effective as, or preferable to IP treatment.	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Liu 2020[17]	CAPS, PCL-S, PHQ-9 For completers, the VTC-C group showed significantly smaller improvement in CAPS scores from baseline to PT (p = 0.004) but showed similar symptom improvements (p = 0.784) from baseline to 6-month follow-up. VTC-C did not show NI in improvement at PT (p = 0.430; difference = 0.58 per week; 95% CI = 0.19, 0.96; NI margin = 0.61), but it showed NI in improvement at 6-month follow-up (p = 0.011; difference = 0.03 per week; 95% CI = -0.17, 0.22; NI margin = 0.26). The ITT analyses showed the same pattern of results. On the PCL, the NI test showed that VTC-C is non-inferior to the IP-C group (p < 0.001; difference = –0.10 per week; 95% CI = –0.37, 0.17; NI margin = 0.40) for both PT and 6-month follow-up improvement (p < 0.001; difference = -0.05 per week; 95% CI = -0.17, 0.07; NI margin = 0.16). In ITT analysis, the NI was still significant for PT change but not for 6-month follow-up (p = 0.070; difference = 0.06 per week; 95% CI = -	Completion of less than 9 sessions Among those who finished baseline visit, 25.6% dropped out at the post-treatment visit. There was no statistically significant difference (p = 0.429) in dropout rates between the IP-C condition (28.2%) and the VTC-C condition (23.1%).	Session attendance During the treatment period in this study, participants in the IP-C group attended an average of 9.7 treatment sessions (SD = 4.0) and participants in the VTC-C group attended an average of 8.9 sessions (SD = 4.5) out of a total 12 sessions. The number of attended sessions was not significantly different between groups (p = 0.153).	Homework completion Percentage of completed homework was 0.47 (SD = 0.29) for the IP-C group and 0.43 (SD = 0.3) for the VTC-C group (p = 0.612).	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
	0.07, 0.20; NI margin = 0.17) when the NI margin is 0.5 SD. There was no significant difference in PHQ–9 score improvement from baseline to PT (p = 0.861) or from baseline to 6-month follow-up (*p* = 0.854) between the VTC-C and IP-C group.				

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Maieritsch 2016[18]	PCL, CAPS, BDI-II The sample size required to declare equivalence was not achieved. A trend was observed of equivalence between treatment arms on the CAPS (δ = -0.5, 95% CI: 12.4–11.4, *p* = 0.094), and a trend was observed of equivalence on the PCL (ratio = 0.92, 95% CI: 0.78–1.09, p = .079). Group differences on depression were not examined.	Failed to receive any intervention, lost to follow up, or discontinued intervention 43.4% (n=20 for VTC-C and n=24 for IP-C group).	None N/A	WAI At PT, Veterans across both conditions were equivalent on therapeutic alliance (ratio = 1.03, 95% CI: 0.98–1.08, *p* < 0.001).	None N/A
Morland 2014[19] & Morland 2011[20]	CAPS No statistically significant differences between groups in CAPS score reductions over any period.	Did not complete at least 10 of 12 sessions 14 participants in the IP-C group and 15 participants in the VTC-C group did not attend at least 10 of 12 sessions. 78.1% in IP-C and 75.4% in VTC-C group completed treatment (*p* = 0.72).	Session attendance, homework completion Sessions attended: IP-C group =10.1 vs VTC-C group = 9.7 (*p* = 0.58); homework completion: IP-C group = 8.8 vs VTC-C group = 8.7 (*p* = 0.79).	CPOSS-VA, Group Therapy Alliance Scale CPOSS-VA: IP-C=63.9 pts vs VTC-C group 66.7 pts (p = 0.36) Group Therapy Alliance Scale: IP-C = 4.5 vs VTC-C = 4.6 (*p* = 0.84).	Adverse event tracking No adverse events were observed for any participant.

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Morland 2015[21]	CAPS The noninferiority hypothesis was supported in that clinical outcomes in the VTC-C condition were not significantly lower than those in the IP-C condition.	Attended less than 10 sessions 34% of randomized participants dropped out. Treatment retention did not differ between modalities.	Session attendance Engagement (mean # sessions) did not differ between modalities.	WAI, CPOSS-VA, Telemedicine Satisfaction and Acceptance Scale At session 2, women in the IP-C group reported higher therapeutic alliance compared to VTC-C (*ES* = 0.07). No differences were found in therapeutic alliance at sessions 6 or 12 or therapists at any time point. At PT, women reported high levels of satisfaction with services. However, there was a statistically significant difference (*t*(345.9) = −2.24, *P* = .03) between treatment modalities with IP-C women reporting scores 4.3 points higher (out of 79) compared to VTC-C women (*ES* = −0.24).	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Morland 2020[22]	CAPS-5, PCL-5, BDI-II Neither a significant main effect for treatment condition nor a significant interaction between treatment condition and time was identified for the CAPS (ps > .385) or the PCL-5 (ps > .086). A significant treatment condition by time interaction was found for BDI-II scores. While there were no differences between IP-H and VTC-C means at pretreatment, compared to IP-H scores, the BDI-II scores in the VTC-C condition were significantly higher at PT (mean = 8.0 points; 95% CI = 0.85, 15.17; *p* = .028) and at 6-month follow-up (mean = 7.3 points; 95% CI = 0.03, 14.61; *p* = .049). Scores for the VTC-H condition did not statistically differ from either the IP-H or the VTC-C conditions.	Completing <8 sessions without meeting definition of early responder Compared to IP-H, participants in both VTC-C (OR = 2.67, 95% CI = 1.10, 6.52; *p* = .031) and VTC-C (OR = 5.08; 95% CI = 2.10, 12.26; *p* < .001) were significantly more likely to drop out.	Session attendance, predictors of dropout Groups differed on the number of sessions completed, F(2) = 4.95, *p* ≤ .01 with participants in the IP-H condition compared to the VTC-C condition attending more PE sessions (9.8 vs 7.0 sessions, *t*(344.7) = 5.54; *p* < .001). Older age of entry into military service predicted a greater risk of dropout, OR = 1.14, 95% CI = 1.01, 1.28, *p* = .036. The average age at entry was 20.0 years (SD = 2.87; range = 17–31 years). None of the other characteristics (including baseline symptoms) predicted risk for dropout, all *p* > .302	None N/A	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Morland 2022[23]	CAPS-5, Brief Inventory of Psychosocial Functioning, and Couples Satisfaction Index There were no significant differences between the IP-C and VTC-H groups. There were no significant changes in PTSD symptoms during the follow-up period for any of the 3 treatment arms, and there were no differences between treatment arms in degree of maintenance of treatment gains. There were no differences in slopes of IPF or CSI scores between treatment groups during treatment or follow-up.	Attended less than 7 out of 8 sessions There were no significant differences in treatment completion rates between IP-C (74%), VTC-H (70%), and psychoeducation control (69%) conditions ($t < .65$, $p > .4$).	None N/A	WAI-short form and CSQ IP-C was not significantly different from VTC-H in working alliance or client satisfaction ($p > .18$).	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Peterson 2022[24]	PCL-5, CAPS-5, BDI-II Improvement on the PCL-5 was about twice as large in the IP-H (*d* = 2.1) and VTC-H (d = 2.0) formats compared to IP-C (d = 1.3). Both of those differences between treatments were statistically large (*d* = .8 and .7) and significant (*p* = .009 and .014). The difference between IP-H and VTC-H PCL-5 outcomes was negligible (*p* = 0.77, *d* = −.08). The differences between treatment arms on the PCL-5 dissipated by the 6-month follow-up point. Differences between treatment arms on the CAPS were not significant (all *p* > .20). Improvement in depression was considerably larger in the IP-H (*d* = 1.2) and VTC-H (*d* = 1.1) arms than IP-C treatment (*d* = .52).	Completion of <75% of sessions A total of 42 patients (35%) dropped out of treatment, including 11 who never began. Differences in dropout between groups were not statistically significant (log-rank X^2 = 2.69, *df* = 2, *p* = .26).	Acceptability The acceptability of the 3 treatment options differed significantly (X^2 = 14.2, *df* = 2, *p* = 0.0008). Among those opting out of 1 delivery modality, most refused IP-H treatment (54%) followed by IP-C (29%) and VTC-H (17%).	None N/A	Adverse events No adverse events differed significantly by group after adjustment for the numbers of participants in each group.
White 2021[25]	PCL-M Treatment modality was not significantly associated with change in PCL scores.	NR NR	None N/A	CPOSS, SDPQ In multivariate regression, treatment modality was a significant predictor of CPOSS (sr^2 = .02, *t* = –6.05, *p* < .001) but not SDPQ.	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Ziemba 2014[26]	CAPS, HAM-A, MADRS, SF-36	NR	None	Patient satisfaction	None
	% Change in mean CAPS score: -24.2% in VTC-C vs -24.4% for IP-C group. % Change in mean SF-36 v2 mental health score: +45.8% for VTC-C and +37.9% for IP-C groups. The HAM-A demonstrated average scores of 34 and 35 for the pretherapy assessments of the VTC-C and IP-C groups, respectively. In turn, postintervention assessment scores for the HAM-A presented a score of 27 each for both groups. The MADRS demonstrated similar pretherapy scores of 32 and 31 for the VTC-C and IP-C groups and respective post-therapy scores of 26 and 23.	NR	N/A	Overall satisfaction scores: 98.1 for VTC-C vs 92.1 for IP-C group.	N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Depression					
Alegria 2014[27] & Aguilera 2018[28] & Alcantara 2016[29] & Kafali 2014[30]	PHQ-9; HSCL-20, WHODAS 2, PSWQ On the HSCL, PHQ-9, and WHODAS there is no significant difference between the impact of the 2 interventions (p = 0.89, 0.69, and 0.91, respectively). Patients in the telephone condition exhibited greater worry reductions than those in the IP-C condition (telephone: M = -7.83, SD = 11.45; IP-C: M = -6.73, SD = 12.23; p = .046).	NR NR	Initiation of care, engagement (total sessions, missed sessions, additional sessions, receipt of prescription for mental health condition, other mental health appointment) Intervention participants in the IP-C group were twice as likely to not initiate care (21.4%) as those in the telephone group (10.3%). There were no significant differences in engagement factors by condition.	Clinician review of homework, homework completion (modified version of the Assignment Compliance Rating Scale), treatment satisfaction, cost-effectiveness These secondary analyses were conducted with a sub-sample (n = 123[28]): Assessed whether the relationship between homework completion and review and outcomes differed based on condition and found no significant differences. Level of satisfaction did not differ between conditions (full sample).[29] Phone CBT was significantly less costly (by $501) and more cost effective than IP-C; for the phone CBT group, one score reduction in PHQ-9 costs $634 less than IP-C (full sample).[30]	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Choi 2014[31]	HAM-D The group differences in predicted mean HAM-D scores were not significant between telephone participants and IP-H participants at 12-week (*t* = -0.31, *df* = 239.61, *p* = 0.755) or 24-week follow-up (*t* = -0.90, *df* = 207.56, *p* = 0.369). Standardized mean difference effect sizes for HAM-D score changes were ES_{sm} = 0.77 for telephone and ES_{sm} = 0.70 for IP at the 12-week follow-up and ES_{sm} = 0.66 for telephone and ES_{sm} = 0.45 for IP-H at the 24-week follow-up.	Completion of less than 6 sessions 14 participants dropped out of the study before completing 6 sessions on IP-H (n = 7), telephone (n = 5), and telephone care calls (n = 2).	None N/A	Treatment Evaluation Inventory (TEI) The telephone group reported slightly higher TEI scores than the IP-H group (72.14 ± 6.64 for telephone and 68.08 ± 8.27 for IP-H; *t* = 2.305, *df* = 70, *p* = 0.024.	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Egede 2015[32] & Egede 2016[33] & Egede 2017[34] & Egede 2018[35]	Geriatric Depression Scale, BDI, SCID, SF-36 Estimated differences in treatment response proportions were 0.88% (90% CI –10·13 to 11·89) for BDI and 2.06% (90% CI –7·46 to 11·58) for GDS for the per protocol sample. The CI lower limits for response proportions are within the allowable NI limit, showing that VTC-H was NI to IP-C at the 12-month timepoint. Differences in treatment response on the SCID (primary outcome) at 12 months were not significant. At PT NI of VTC-H could not be established because the lower limit of the 90% CI was not less than Δ and superiority of IP-C could also not be established because 0 is contained in the CI (for per protocol sample). ITT for 8 weeks (PT): -7.39% (-17.34 to 2.56) for the BDI, -7.79% (-17.01 to 1.43) for the GDS. ITT for 3-months: -3.72% (-13.86 to 6.41) for the BDI, -3.01 (-11.55 to 5.52) for the GDS. None of the SF-36 scores show a significant difference	Did not complete all 8 sessions 192 (80%) of participants completed all 8 sessions; 97 (81%) in the VTC-H group vs 95 (79%) in IP-C.	None N/A	CPOSS, treatment credibility, SDPQ, health care cost trajectories, cost-effectiveness There was no statistical difference in patient satisfaction at any time point. For service delivery perception, there was no statistical difference for most of the variables. The IP-C group did produce statistically superior scores at study end on likelihood for return and overall satisfaction. The results became nonsignificant following a Bonferroni correction. Health care costs before, during, and after the intervention did not differ between groups (treatment completers only). Although the intervention costs for VTC-H were higher, Veterans receiving BA via telehealth had lower health utilization costs 1 year after the intervention than those receiving IP-C care while	Adverse events Did not note any adverse events for any participant in the study.

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
	between the 2 treatment groups.			the QALYs were approximately the same.	
Hungerbuehler 2016[36]	HAM-D, MHI-38, medication course, relapse The VTC-H group showed significantly higher levels of depression (mean IP-C score: 6.19 (3.61); mean VTC-H score: 7.92 (3.59); *P* = .01) and lower levels regarding mental health status (mean IP-C score: 132.89 (25.49); mean VTC-H score: 121.25 (26.19); *P* = .02) than the IP-C group at baseline. At 6 and 12 months, the initial group differences with respect to severity of depression and mental health status were no longer significant. No group differences were found with respect to the type and dosage of medication at the 6- and 12-month follow-up. Five participants were excluded because they relapsed (scored higher than 17 on the HAM-D); 4 in the IP-C and 1 in the VTC-H group.	'Lost to follow-up' included 'discontinued intervention' and 'other reasons.' Discontinuation of treatment occurred if participants were fully remitted, missed 3 consultations in a row, had a relapse (HAM-D > 17), needed additional care, or showed an elevated suicide risk. At 6 months, there were significantly more dropouts in the IP-C group (n=10) than in the VTC-H group (n=3; X^2_1 = 4.143, *P* = .04). At 12 months, there were still more dropouts in the IP-C group (n=14) than in the VTC-H group (n=8), but the difference was no longer significant.	Treatment adherence, medication adherence (MMAS-4) Participants in the IP-C group tended to miss more appointments than participants in the VTC-H group (F_{105} = 0.753, *P* = .06). There were no significant group differences regarding medication compliance at 6 and 12 months between the 2 groups. Participants in the IP-C group tended to be more adherent than participants in the VTC-H group at 12 months (X^2_1 = 2.864, *P* = .07).	CSQ, WAI-S There were no significant differences between treatment conditions with respect to satisfaction of the participant at 6 and 12 months. There were no group differences with respect to working alliance at either of the follow-ups.	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Luxton 2016[37] & Bounthavong 2018[38] & Pruitt 2019[39] & Smolenski 2017[40]	BHS, BDI-II, BAI, PCL-M At PT participants in the IP-C group had an average reduction of 6.21 points on the BHS (95% CI 7.38, 5.05) and 17.63 points on the BDI-II (95% CI 20.21, 15.06), while participants in the VTC-H group had an average reduction of 3.91 points on the BHS (95% CI 5.25, 2.57) and 13.40 points on the BDI-II (95% CI 16.36, 10.44). For both outcomes, the magnitude of decrease over time was less pronounced for the VTC-C group compared to the IP-C group. The upper bound of the 90% confidence interval included 0.50. The per protocol analysis provided similar results. There were no statistically significant differences between the treatment groups on anxiety and PTSD outcomes.	Did not complete all 8 sessions The difference in proportions of subjects that did not complete treatment between the groups was not statistically significant (VTC-H = 35.48%, IP-C = 28.81%, X^2= 0.62, df = 1, p = .433).	None N/A	CSQ, cost Patient satisfaction was very high, with no significant differences between the treatment modalities. Higher end-of-treatment satisfaction for IP-C care was most commonly associated with younger age and more junior military status. Conversely, higher end-of-treatment satisfaction for VTC-H was more commonly associated with older, more senior, and more symptomatic service members. In the base case analysis the total direct cost of VTC-H was higher than IP-C (US $71,974 vs US $20,322). Assuming that patients possessed government-approved video-conferencing technology, VTC-H was less costly compared to IP-C care (US $19,177 vs US $20,322).	Treatment session checklist There were 7 participants who had adverse events that required reporting in the VTC-H group and 4 in the IP-C group. None of these adverse events were determined to be related to study procedures. The safety protocol was initiated one time; this occurred for a military service member in the VTC-H condition who contacted his study provider and presented to the research staff.

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Mohr 2012[41] & Kalapatapu 2014[42] & Stiles-Shields 2014[43] & Stiles-Shields 2015[44]	HAM-D, PHQ-9, AUDIT Post-treatment avg symptom change: IP-C group −10.32 (95% CI −9.02, −11.62) points on the HAM-D and −10.03 (95% CI −11.05, −9.00) on the PHQ-9 Telephone group −9.25 (95% CI −10.42, −8.09) points on the HAM-D and −10.12 (95% CI 11.08 −9.15) on the PHQ-9 6-mo follow up average change in symptom severity IP-C group −10.69 (95% CI −11.99, −9.39) points on the HAM-D and −10.46 (95% CI −11.53, −9.39) points on the PHQ-9 Telephone group −7.78 (95% CI −8.98, −6.57) points on the HAM-D and −8.35 (95% CI −9.40, −7.29) points on the PHQ-9 Sub-group (n=103) with co-occurring AUD: The decrease in AUDIT total score from baseline to end of treatment was significant in participants in	Discontinuation of treatment before 5 or 18 sessions Attrition before 18 sessions: lower in telephone (n = 34; 20.9%) than in IP-C (n=53; 32.7%; *P* = .02). Attrition before 5 sessions: lower in telephone (n=7; 4.3%) than in IP-C (n = 21; 13.0%; *P* = .006). Sub-group with co-occurring alcohol use disorder: groups did not significantly differ in failure to engage in treatment, failure to complete treatment, or discontinuation of treatment.	Treatment adherence (number of sessions completed) Telephone patients attended significantly more sessions (mean, 15.5; median, 17; SD, 4.4; interquartile range, 16-18) than those receiving IP-C CBT (mean [SD], 13.7 [6.1]; median [IQR],17 [11-18]; *P* = .003). Sub-group with co-occurring alcohol use disorder: groups did not significantly differ in number of sessions attended.	WAI There were no significant differences in client or therapist WAI between telephone or IP-C (Cohen's f 2 ranged from 0 to .013, all *ps* > .05).	Adverse event tracking There were no adverse events (*eg*, suicide, suicide attempt, psychiatric hospitalization) for either treatment condition.

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
	each group ($P < 0.001$); this decrease in AUDIT total score was not significant between groups ($P = 0.21$). Prediction models: Treatment delivery method did not impact the prediction of outcome by baseline demographics and symptom severity.				

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Ruskin 2004[45] & Rhein 2001[46]	HAM-D, BDI, Spielberger State Anxiety Scale, GAF, CGI Depressive symptoms (HAM-D) significantly improved over the treatment period (time main effect: F = 49.0, df = 3, 117, p < 0.001), and improvement did not differ by treatment group (time-by-treatment interaction: F = 0.4, df = 3, 117, n.s.). BDI, state anxiety scale, GAF, CGI, and Short-Form Health Survey scores improved significantly over the course of treatment, and improvement did not differ significantly between the 2 treatment groups. Results indicated that there was no significant interaction between treatment type and personality disorder measured dichotomously or continuously (subset of first 90 patients randomized).	Unwilling to continue with sessions + no final session to assess final symptom severity 16 participants (27%) in the VTC-C group and 18 (30%) in the IP-C group dropped out of the study (X^2 = 0.4, df =1, n.s.). A time-to-event analysis indicated no significant difference in the time course of the dropouts between the 2 treatment groups (X^2 = 0.1, df = 1, n.s.).	Session attendance, medication adherence (taking at least 70% of prescribed doses). Both groups kept appointments for an average of 6.5 visits during the study period (t = 0.2, df = 117, n.s.) There was no difference in the percentage of adherent patients between the 2 treatment groups (X^2 = 0.2, df = 1, n.s.).	Ad-hoc patient and psychiatrist satisfaction scales & resource consumption and cost effects There was no difference in patient satisfaction between groups at visit 4 (t<0.1, df=87, n.s.), visit 6 (t=–0.4, df=74, n.s.), or visit 8 (t=1.3, df=74, n.s.). Psychiatrist satisfaction was greater for IP-C (t = -2.2, df=79, p < 0.05). The estimated marginal costs to the institution were $86.16 for a VTC-C session and $63.25 for an IP-C treatment session (t = 3.2, p < 0.001). However, when the cost of psychiatrist travel time was factored in and the time-distance effect was modeled, the cost of remote treatment was equal to that of in-person treatment if the psychiatrist had to travel 22 miles from the medical center to the	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
				clinic and was less if the psychiatrist had to travel more than 22 miles to the clinic. The VTC-C group was not associated with significantly different overall consumption of Veterans Health Administration health care (t = 0.7, ns).	

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Anxiety					
Cherestal 2019[74]	Flight related anxiety (FAS, FAM), behavioral avoidance (graduation 'flight' at end of treatment) Statistically significant differences were not observed between the telephone and IP-C conditions on the final total FAS scores [F(1, 17) = 1.13, *p* = .31], or anticipatory [F(1, 17) = 5.30, *p* = .48], in flight anxiety [F(1, 17) = 0.39, *p* = .55], and general anxiety subscales [F(1, 17) = 3.67, *p* = .08] after controlling for initial FAS score. There were also no statistically significant differences between the final total FAM score [F(1, 17) = 1.00, p = .34] or the somatic [F(1, 17) = 1.83, p = .20] and cognitive subscales [F(1, 17) = 0.04, *p* = .95] between treatment conditions. At the conclusion of treatment, participants did not significantly differ in rates of completion of graduation flights as a function of treatment condition (*p* = 1.00).	Did not complete treatment 4 participants of out 22 randomized did not complete treatment	None N/A	Presence in the virtual environment (PQ), WAI Total scores on the PQ showed a non-significant pattern of being higher in the IP-C condition than the telephone condition. Differences in scores between the total scores on the final administration of the WAI between treatment conditions did not reach statistical significance [F(1, 17) = 4.37, *p* = .06], after controlling for initial FAS score. There was a statistically significant difference between the final scores of the task subscale [F(1, 17) = 6.46, *p* = .03] and the goal subscale [F(1, 17) = 5.52, *p* = .04] between treatment conditions.	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Lovell 2006[48]	Y-BOCS, BDI Clinical outcome at all 4 time points was equivalent for both treatment arms. At 6 months of follow-up the adjusted estimate of the effect of treatment was 0.70 (− 2.71 to 4.11) for the Y-BOCS and 1.51 (− 2.23 to 5.25) for the BDI - a slight reduction in the mean value for telephone compared with IP-C delivery. All confidence intervals for the Y-BOCS score are within 5 units; this suggests that the treatments are equivalent. Treatment was deemed clinically relevant if the mean pretreatment score was reduced by 2 SDs or more after treatment. Treatment was clinically relevant in 49 patients (72%)—27 (77%) patients in the telephone arm and 22 (67%) in the IP-C arm.	Did not complete the intervention 3 patients in the IP-C group and 1 patient in the telephone group did not complete the intervention	None N/A	CSQ Patients were very satisfied with treatment, and the results were similar for both treatments.	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Watts 2020[49]	ADIS-IV, SCID-II The results revealed a significant clinical and statistical improvement of the participants between the beginning and end of the therapy. There was no significant difference between the 2 treatment conditions.	Did not complete treatment 14/79 allocated to IP-C and 17/69 allocated to VTC-C did not complete treatment.	None N/A	WAI Clients assessed working alliance significantly higher in the VTC-C group compared to IP-C, t(113) = -2.75, p = .007, η^2 .06. There was no significant difference in psychotherapists' assessment of the quality of the working alliance according to treatment condition, t(113) = .158, p = .89.	Adverse event tracking No adverse events were reported by participants, psychotherapists, or independent evaluators rating treatment fidelity.
Multiple mental health conditions					
De Las Cuevas 2006[50]	SCL-90R, CGI There was no statistically significant difference in the mean change in CGI-Improvement and SCL-90R Global Indexes scores between groups.	Did not complete treatment Of 140 randomized patients, 130 completed 24 weeks of treatment. Only 4 patients dropped out prematurely from the VTC-C group and 6 from the IP-C group.	None N/A	None N/A	None N/A

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
O'Reilly 2007[51]	BSI, SF-36	NR	None	CSQ, cost	Hospitalization
	All results support the hypothesis that telepsychiatry produces equivalent outcomes to face-to-face care.	NR	N/A	The CSQ indicated a moderate degree of satisfaction and analysis showed equivalence between the 2 groups. The average cost of VTC-C was 10% less per patient (16% less per visit) than the cost of IP-C service.	Conducted an ITT analysis for the proportion of participants hospitalized within 12 months from initial consultation, and this analysis showed equivalence between the 2 groups.

Author, Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Stubbings 2013[52]	DASS, QLES, BDI-II, Obsessive-Compulsive Inventory, Health Anxiety Questionnaire, PSWQ, and/or the Anxiety Sensitivity Index as needed There was no significant main effect for condition on any of the DASS subscales (Depression $F_{1,58} = 1.98$, $P = .16$; Anxiety $F_{1,58} = 0.69$, $P = .41$ and Stress $F_{1,58} = 2.11$, $P = .15$) or the QLES ($F_{1,62} = 0.25$, $P = .62$). The effect sizes ranged from small to medium (d=0.37, 0.22, 0.38, and 0.13 respectively).	Attended less than 8 sessions 1/12 IP-C participants dropped out vs 1/14 VTC-C participants	None N/A	WAI-S, CSQ There were no significant differences between conditions in client ratings of the WAI-S ($t_{21} = –0.63$, $P = .53$, one-tailed, $d = –0.26$), or in therapist ratings ($t_{22} = 0.53$, P =. 60, one-tailed, $d = 0.23$), and client ratings of the CSQ ($t_{22} = –0.29$, $P = .77$, one-tailed, $d = –0.12$).	None N/A

Abbreviations. ADIS-IV=Anxiety Disorders Interview Schedule; AUD=alcohol use disorder; AUDIT=Alcohol Use Disorders Identification Test; BAI=Beck Anxiety Inventory; BDI-II=Beck Depression Inventory-Second Edition; BHS=Beck Hopelessness Scale; BSI=Brief Symptom Inventory; CAPS=Clinician Administered PTSD Scale; CBT=cognitive behavioral therapy; CGI=Clinical Global Impressions Scale; CI=confidence interval; CPOSS=Charleston Psychiatric Outpatient Satisfaction Scale; CSQ=Client Satisfaction Questionnaire; DASS=Depression Anxiety Stress Scales; FAM-Flight Anxiety Modality; FAS=Flight Anxiety Situations; GAF=Global Assessment of Functioning; GDS=Geriatric Depression Scale; GSI=Global Severity Index; HAM-A=Hamilton Anxiety Rating Scale; HAM-D=Hamilton Depression Rating Scale; HSCL-20=20-Item Hopkins Symptom Checklist; IP-C=clinic-based in-person; IP-H=home-based in-person; IPF=Inventory of Psychosocial Functioning; ITT=Intent-to-treat; MADRS=Montgomery-Asberg Depression Rating Scale; MHI-38=Mental Health Inventory; MMAS-4=Morisky Medication Adherence Scale; NI=noninferiority; NR=not reported; OBT=office-based telehealth; OR=odds ratio; PCL-5=PTSD Checklist for DSM-5; PCL-M=PTSD Checklist-Military Version; PDS=Posttraumatic Stress Diagnostic Scale; PE=prolonged exposure; PHQ-9=Patient Health Questionnaire-9, PQ=Presence Questionnaire; PSWQ=Penn State Worry Questionnaire; PT=post-treatment; PTSD=post-traumatic stress disorder; QLES=Quality of Life Enjoyment and Satisfaction Questionnaire; QOLI=Quality of Life Inventory; SCID-II=Structured Clinical Interview for DSM-IV Axis I Disorders; SCL-90-R=Symptom Checklist 90 Revised; SF-36=Short Form-36; SD=standard deviation; SDPQ=Service Delivery Perception Questionnaire; SE=standard error; TAU=treatment as usual; TCS=Treatment Credibility Scale; TH=telehealth; VTC-C=clinic-based videoteleconference; VTC-H=home-based videoteleconference; WAI=Working Alliance Inventory; WAI-S=Working Alliance Inventory – Short Version; WHODAS-2=WHO Disability Assessment Schedule; Y-BOCS=Yale-Brown Obsessive Compulsive Scale.

Cohort Studies

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Posttraumatic stress disorder					
Germain 2009[53] & Germain 2010[54]	SCID, m-PSS, BDI, BAI, ACF 81% of the participants assigned to VTC-C and 75% of participants in the IP-C condition no longer met diagnostic criteria for PTSD after treatment. Repeated measures ANOVAs completed for the overall m-PSS score do not show a significant interaction effect, (f = 1.14, η2 = 0.01). Similar analyses were done on all the secondary measures (BDI [f = 1.09, η2 = 0.02], BAI [f = 0.46, η2 = 0.01], and ACF [f = 1.01, η2 = 0.02]; they did not reveal a significant interaction or group effect.	Did not complete therapy 8 participants dropped out of the VTC-C treatment and 12 dropped out of IP-C treatment. In addition, 6 individuals in the VTC-C condition and 7 in the IP-C condition were excluded from therapy.	Session attendance Participants in the VTC-C condition received an average of 21 sessions and those in the IP-C condition an average of 19 sessions.	WAI, SEQ, DCCS, VTS, VT-Q and VT Sessions-Q There was no significant difference in alliance development between conditions. For the SEQ, there was no significant difference between correlations in the 2 treatment conditions ($p > 0.05$ for all 5 subscales). Partial correlations were calculated to explore the relationship between the effectiveness of therapy and the different variables associated with the use of VTC. The results did not reveal any significant correlation between the various questionnaires the overall posttreatment m-PSS score.	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Gros 2011[55]	PCL-M, BDI-II There was a significant effect of setting on the posttreatment PCL-M and BDI-II after controlling for pretreatment scores (F > 13.4, p < .01, d > .56), suggesting that patients receiving IP-C treatment evidenced greater symptom reductions than VTC-C on both measures.	Patients who finished 12 sessions but did not complete treatment 24/62 participants in VTC-C did not complete treatment. The IP-C comparison group was limited to treatment completers.	Session attendance, predictors of dropout The VTC-C (M = 11.0, SD = 1.9) and IP-C (M = 9.2, SD = 2.4) groups varied in their number of sessions completed, F(1, 63) = 10.7, p < .01. There were no differences in any of the baseline measures of symptomatology between the completers and the noncompleters (F < 2.5, p >.05)	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Knowlton 2021[56]	PCL-5, BDI-II The magnitude of changes in PTSD and depression symptoms was consistent across treatment delivery modalities when pooling PE and CPT into a single variable, as measured by the Reliable Change Index. Within condition effects were assessed using a 2-way ANOVA for posttreatment scores by treatment type and method of delivery. The analyses were not significant for PCL-5 scores, F(2, 272) = 0.19, *p* = .83, *d* = .20, or BDI-II scores, F(2, 272) = .17, *p* = .91, *d* = .17.	Ending treatment prior to optimal benefit (defined as 8 sessions) The VTC-H group had the highest treatment completion rate (56%), followed by the IP-C (50%), and VTC-C (46%) groups. There were no significant differences across treatment modality in terms of completion of minimum required dosage χ^2(2, N = 274) = 6.6, *p* = .19, and attending up to 7 sessions, χ^2(2,N = 137) = 2.5, *p* = .34, *d* = .16.	Session attendance, homework completion There were no significant differences across treatment modality in terms of completion of minimum required dosage X^2(2, N = 274) = 6.6, *p* = .19, and attending up to 7 sessions, X^2(2,N = 137) = 2.5, *p* = .34, *d* = .16. Average number of sessions attended for the IP-C group was 7.5, followed by 6.2 for VTC-C, and 8.9 for VTC-H. There were no significant differences in homework completion by method of treatment delivery *t*(273) = 1.21, *p* = .98, *d* = .17	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
LoSavio 2021[57]	PCL-5 In a propensity score-matched analysis, which resulted in 103 cases per group, there was no significant time by treatment format interaction (*b* = -.073, *t*(101.37) = -1.07, *p* = .29)	Completion of less than 5 sessions Patients receiving care via VTC were significantly more likely to complete treatment [X^2(1, N = 228) = 3.97, *p* = .046] than patients receiving care IP (34.0% dropout rate for IP vs 21.3% for VTC)	None N/A	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Tuerk 2010[58]	PCL, BDI-II Mean pre- and posttreatment PCL scores for the IP-C group were 60.7 (*SD* = 9.5) and 27.7 (*SD* = 6.0), respectively. This difference is clinically and statistically significant, *t*(28) = 16.9, *p* < .001, *d* = 4.2. Mean pre- and posttreatment PCL scores for the VTC-C group were 61.0 (*SD* = 10.6) and 34.9 (*SD* = 7.6). This difference is also clinically and statistically significant, *t*(8) = 12.3, *p* < .001, *d* = 2.9. Mean pre- and posttreatment BDI-II scores for the IP-C group were 27.8 (*SD* = 9.3) and 10.9 (*SD* = 6.4). This difference is clinically and statistically significant, *t*(28) = 8.7, *p* < .001, *d* = 2.2. Mean pre- and posttreatment BDI-II scores for the VTC-C group were 27.3 (*SD* = 12.5) and 7.6 (*SD* = 4.7). This difference is also clinically and statistically significant, *t*(8)=5.5, *p* < .001, *d* = 2.3. Potential differences between groups were not modeled due to low power.	Completed <5 sessions The treatment completion rate for the IP-C sample was 83% and the treatment completion rate for the VTC-C sample was 75%.	Sessions attended The mean number of sessions for treatment completers was 10.1 (*SD* = 3.8) for IP-C group and 10.0 (*SD* = 6.3) for the VTC-C group.	None N/A	Adverse events There were no instances in the VTC-C condition where on-site staff had to be contacted for patient safety.

Author Year	**Mental Health Symptom Measures** **Mental Health Symptom Results**	**Dropout Definition** **Study Attrition**	**Access and Continuity of Care Measures** **Access and Continuity of Care Results**	**Quality and Implementation-related Measures** **Quality and Implementation-related Results**	**Harm Measures** **Harm Results**
Valentine 2020[59]	None N/A	Completion of less than 12 sessions or completion of less than 8 sessions (minimally adequate care) When attrition was assessed using a 12-session period, completion rates were 50% and 32.3% for IP-C and VTC-H, respectively. This difference approached but did not reach statistical significance ($X^2(1) = 3.21$, $p = 0.073$. When defining treatment completion as 8 or more sessions, 62.1% of individuals completed treatment delivered IP-C vs 41.9% of individuals completing treatment delivered via VTC-H.	None N/A	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Wierwille 2016[60]	PCL-S, BDI-II Although PCL-S scores decreased significantly from pre- to posttreatment for both groups, there was a greater decrease in scores in the IP-C group than in the VTC-C group. Though BDI-II scores decreased significantly from pre- to posttreatment for both groups, there was a greater decrease in scores in the IP-C than in the VTC-C group.	Appears to be termination of treatment without documented agreement between the patient and provider to terminate treatment because the full therapy protocol had been completed, sufficient symptom alleviation had occurred, or no symptom alleviation had occurred despite a majority of the protocol having been completed. Though the VTC-C group had a 13% numerically higher dropout rate than the IP-C group (VTC-C = 60.00%, IP-C = 47.06%), $X^2(1) = 3.51$, $p = .061$, this difference did not reach statistical significance.	None N/A	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Depression					
Ritchie 2007[60]	BDI-II Without respect for session number, the BDI-II scores of the group conducted via VTC-C (*M* = 21.58) did not differ from that of the group conducted IP-C (*M* = 20.13). No significant interaction between intervention and time/session (*F*(1, 8) = .19, ns).	NR NR	Session attendance The participants in the VTC-C group had a 98.21%compliance rate. The participants in the IP-C group had a 71.42% compliance rate. It is suspected that the difference in compliance rate was due to the status of the participants in each group and the degree of effort required for participants to get to the therapy session location.	Group Climate Questionnaire Participants from the VTC-C group reported greater GCQ-S engaged scores, (*M* = 3.63), than did participants from the IP-C group, (*M* = 2.65). No significant interaction between time/session and intervention group with respect to GCQ-S engaged scores *F*(1, 8) = .11, ns. No observed main effect for therapy group on GCQ-S conflict subscale, *F*(1, 8) = .11, ns. However, participants from the group conducted via VTC-C significantly increased their report of conflict from session one to session 8, *t*(8) = -2.33, *p* < .05, but participants in the IP-C group did not, *t*(8) = 1.09, ns. No main effects were found for GCQ-S	NR

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
				avoiding subscale; therapy group, *F*(1, 18) = 1.72, ns; session, *F*(1, 8) =.48, ns. These 2 null effects were followed by a null for the test of the interaction between group and session, *F*(1, 8) = 66, ns.	

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Anxiety					
Bouchard 2004[62]	Panic attacks, panic apprehension (daily diary); ACQ, BSQ, Mobility Inventory, Self-Efficacy to Control a Panic Attack Scale, STAI, BDI, SDS At PT, 90% of the participants in the IP-C condition were panic free, as were 81% of those in the VTC-C condition. At follow-up, 100% of the participants in the IP-C condition were panic-free, as were 91% of those in the VTC-C condition (all chi-squared values were not significant). There was only 1 significant group–time interaction over the pre–post interval that met the a priori fixed level of significance. The VTC-C group had a greater reduction in panic frequency than did the IP-C group. Although significant at p = 0.05, the interaction for panic apprehension did not reach the 0.025 level of significance, which was fixed a priori.	NR NR	None N/A	WAI, treatment credibility Alliance scores are reported only for participants in the VTC-C condition and were measured after the 1st session, after the 3rd session, and at PT. The working alliance scores of the participants in the VTC-C group were very high at each assessment: 222.2 after the first session, 243.1 after the fifth session, and 242.4 at posttreatment.	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Bouchard 2020[63]	Panic and Agoraphobia Scale, Mobility Inventory, ACQ, BSQ, BDI Contrasts for pre to post by condition interaction were nonsignificant and trivial for all measures, except for the fear of body sensations which was very small [for PAS (F(1,69) = 0.2, *p* = 0.63, η_p^2 = 0.003), for MI (F(1,69) = 0.08, *p* = 0.78, η_p^2 = 0.001), for BSQ (F(1,69) = 1.65, *p* = 0.2, η_p^2 = 0.023), and for BDI (F(1,69) = 0.098, p = 0.76, η_p^2 = 0.001)]. Gains were all maintained at the 12-mo follow-up. PT to follow-up contrasts were nonsignificant [for PAS (F(1,69) = 1.97, *p* = 0.17, η_p^2 = 0.028), for MI (F(1,69) = 0.02, *p* = 0.87, η_p^2 = 0.000), for BSQ (F(1,69) = 3.32, *p* = 0.07, η_p^2 = 0.046), and for BDI (F(1,69) = 0.007, *p* = 0.93, η_p^2 = 0.000)]. The non-inferiority tests revealed that VTC-C was statistically no less effective than IP-C on the primary outcome variable, and 2 of the 3 secondary outcome measures (agoraphobia and depressed mood). However, the noninferiority test did not reach statistical	NR A chi-square analysis was conducted to identify differences in dropout rates between VTC-C and IP-C and the result was not significant [X^2(1) = 0.06, ns].	None N/A	WAI, CALPAS, CMOTS The 2 measures of working alliance were administered after sessions 1, 5, and 12. A significant time effect was found with each measure, while no condition or time x condition effects were statistically significant. Motivation toward therapy was high and self-determined in participants in the VTC-C (Mean = 13.26, *SD* = 3.77) and the IP-C (Mean = 12.30, *SD* = 4.65) conditions. The difference in motivation across conditions at pre-treatment was not significant [*t*(67) = 0.96, *p* = 0.34; η_p^2 = 0.01].	Adverse effects No adverse effects were reported.

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
	significance for the fear of body sensations.				

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Milosevic 2022[64]	DASS, SPIN, PDSS-SR, PWSQ, OCI-R There was a significant effect of group in the full-sample analysis, *b* = .05, *SE* = .01, *t* = 4.74, *P* < 0.001, *d* = .23, and in the GAD subsample, *b* = .69, *SE* = .01, *t* = 2.70, *P* = 0.008, *d* = .20. These results indicate a small but significant positive effect of IP-C treatment on reduction in symptom severity over time, relative to VTC-H treatment. This effect was not significant for SAD (*P* = 0.11), PDA (*P* = 0.17), and OCD (*P* = 0.69) groups. Both group, *b* = .03, SE = .01, *t* = 3.48, *P* < 0.001, *d* = .19, and baseline z-score, *b* = -.04, SE = .01, *t* = -6.97, P < 0.001, *d* =-.37, were independently significant predictors of change in symptom severity over time, but the interaction was not, *b* = .01, SE = .01, t = .53, *P* = 0.60, *d* = .03. These results indicate that participants who received F2F treatment achieved greater symptom severity reductions regardless of baseline severity.	Attending <8 sessions and not attending the final 3 sessions Treatment dropout did not differ significantly between groups, with observed dropout rates of 24.5% IP-C and 15.3% for VTC-H, $X^2(1, N = 371) = 3.18$, *P* = 0.08, ϕ= .09.	Session attendance Significantly more sessions were attended by participants in the VTC-H vs IP-C GAD groups, t(146) = -2.02, *p* = .045, *d* = .17. The difference in attendance between the delivery formats approached significance for the full sample, t(371)=-1.94, *p* = .053, *d* = .10.	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Pinciotti 2022[65]	Length of stay, QLESQ, QIDS, Y-BOCS-SR There was a significant difference in length of stay comparing IP-C and VTC-H groups (t[465] = -2.51, p < .05), such that the IP-C group stayed on average 23.22 days and the VTC-H group stayed on average of 25.79 days. There were no significant differences comparing PHP and IOP groups at discharge for any of the YBOCS-SR, QIDS, or QLESQ. There was no significant difference in treatment response between IP-C and VTC-H groups, where 37.61% of the IP-C group fit this criterion and 34.62% of the VTC-H group fit this criterion.	NR NR	None N/A	None N/A	None N/A
Substance use					
Bean 2022[66]	DASS No significant time x group interaction was found for depression scores, F(1, 67) = 0.006, p = 0.94, η_p^2 <0.001, anxiety, F(1, 67) = 0.016, p = 0.90, η_p^2 <0.001, or stress, F(1, 67) = 0.059, p = 0.81, η_p^2 <0.001.	NR NR	Session attendance There were no significant differences observed in the number of sessions attended t(67) = 0.12, p = 0.90 or number of sessions missed for any reason t(67) = 0.51, p = 0.61.	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Vakkalanka 2022[67]	None N/A	Treatment gap of 14 days or more The risk of discontinuation among those with a documented VTC encounter was 0.69 (95% CI: 0.60, 0.78) times that of someone with only an IP SUD encounter. The results of VTC for mental health encounters were similar to those observed for SUD (aHR: 0.69; 95% CI: 0.62, 0.76). Discontinuation appeared to be lower among those with VTC only compared to IP only for both SUD (aHR: 0.48, 95%CI: 0.37, 0.62) and mental health encounters (aHR: 0.46; 95%CI: 0.33, 0.65).	None N/A	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Zheng 2017[68]	Abstinence, time to 30 and 90days abstinence The telepsychiatry group percentage of those attaining 90 consecutive days of abstinence is 49%, and 37% in the IP-C group; Chi-squared test (p = 0.31) indicating no significant difference between the 2 groups. More than half of each group (51% of the VTC-C and 63% of the IP-C group) was unable to attain 90 days of abstinence. Mean/median time to 30 days abstinence: 35/30 days VTC-C vs 42/30 days IP-C (p=0.09) Mean/median time to 90 days abstinence: 106/90 days VTC-C vs 112/94 days IP-C (p=0.22)	Unclear 6 (13%) patients from the VTC-C group and 4 (7%) patients from the IP-C group dropped out of program (mostly had no shows to groups) at an early treatment phase (some less than 1 week), before a possible relapse could be recorded.	Treatment retention For patients who could have potentially stayed in treatment for 90 days, both groups retained close to 50% of patients at 90 days. The retention rates at 365 days decreased to 10/24 (41.7% VTC-C group) and 11/31 (35.5% IP-C group). This difference was not statistically significant between the 2 groups (*p* = 0.99).	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Multiple mental health conditions					
Gannon 2021[69]	PHQ-9, GAD-7 Comparisons of score change on GAD-7 (t = 0.14, p = ns) and PHQ-9 (t = 0.08, p = ns) were not significantly different between the 2 time periods. However, an examination of GAD-7 and PHQ-9 improvement based on categorical change scores showed that significantly more patients had PHQ-9 improvement in the pre-COVID-19 period (46.8%) than in the period during COVID-19 (37.0%; X^2 = 3.834, p = 0.05)). No difference in categorically determined improvement between time periods was observed for GAD-7 (X^2 = 0.75, p = ns). Although there were no overall differences in score changes for the 2 measures based on diagnosis, patients with anxiety spectrum disorders in the COVID-19 period had significantly higher improvements in mean (F = 4.49, p = 0.004) and categorically determined changes (X^2 = 9.15, df = 3, p = 0.027) in GAD-7	NR NR	Patient participation (weeks of enrollment, number of assessments completed) The average number of weeks of patient enrollment in the pre-COVID-19 time period was 3.9 ± 7.6, which was not statistically different from during the COVID-19 time period (4.9 ± 8.9 weeks). Similarly, the mean number of GAD-7 and PHQ-9 assessments completed per patients was statistically equivalent between time periods (3.25 ± 2.3 in the pre-COVID-19 and 2.84 ± 2.4 in the COVID-19 time periods).	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Hammond 2012[70]	PHQ-9, GAD-7, WSAS An unadjusted comparison indicated significant differences between treatments in the reductions of PHQ-9 and GAD-7 symptom scores; telephone intervention appeared to be more effective (PHQ-9: F = 17.5, *p* < .001, effect size (ES)= 0.13; GAD-7: F = 5.93, p= 0.015, ES = 0.07; WSAS: F= 2.82, *p* = 0.087, ES = 0.04). Significant differences between treatment groups on both symptom measures were still observed after controlling for number of assessments, provider sites, and baseline symptom severity measures (PHQ-9: F = 10.9, *p* < .001, ES = 0.14, GAD-7: F= 8.13, *p* = 0.042, ES = 0.10, WSAS: F = 3.12, *p* = 0.078, ES = 0.03). Non-inferiority in favor of telephone treatment for symptom severity persisted as small to moderate effects for all but individuals with the highest symptom severity. In the most stringent comparison, the one-to-one propensity matching, adjusted mean differences in treatment outcomes indicated non-inferiority between telephone vs IP-C	N/A (only included treatment completers) N/A	None N/A	Cost The per-session cost for telephone was 36.2% lower than IP-C. Under the assumption that telephone costs 32.6% (approx. 1.5 times) less than IP-C, the cost per session was estimated for a base-case scenario. The mean cost per session for telephone was £79.19 (95% CI 55.0 to 103.3) and IP-C was £118.76 (95% CI 82.5 to 155.0). Even when the cost ratio was varied from 1.2 times to 2 times, telephone was still cost-effective.	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
	treatments for PHQ-9 and GAD-7, whereas the evidence was moderate for WSAS.				
Khatri 2014[71]	BDI-II According to the change in pre–post score on the BDI-II, similar proportions of participants in the IP-C and VTC-H groups demonstrated positive, negative, or no change in severity of symptoms. Sixty percent of participants in each group showed improvement in their BDI-II score severity (*eg*, pre–post change from a moderate to low BDI-II score), with 20% moving to a more severe category and 20% showing no change in category.	Withdrawals (participants accepted into the program who chose to withdraw) 3 participants withdrew (one in the IP-C group and 2 in the VTC-H group) from the intervention program for reasons unrelated to the study intervention.	None N/A	Group session themes (qualitative analysis). Qualitative theme analysis was conducted on sessions 3, 4, 7, 8, 12, and 13. The themes generated were similar between the 2 group intervention formats.	Technical glitches Technical glitches due to problems with the video conferencing software were frustrating for both the therapist and the study participants. For example, some subjects had difficulty logging on for a session, and the sound was occasionally disrupted during sessions.
Liou 2022[72]	PHQ-9, GAD-7 The adjusted mean differences of change in PHQ-9 and GAD-7 scores between those who underwent virtual EMDR and those treated IP-C were not statistically significant at 0.26 (CI: -2.66 to 3.18) and 0.16 (CI: -2.34 to 2.67), respectively.	N/A N/A	None N/A	None N/A	None N/A

Author Year	Mental Health Symptom Measures Mental Health Symptom Results	Dropout Definition Study Attrition	Access and Continuity of Care Measures Access and Continuity of Care Results	Quality and Implementation-related Measures Quality and Implementation-related Results	Harm Measures Harm Results
Zimmerman 2017[73]	RDQ-M Scores on the RDQ-M subscales at admission and change scores from admission to discharge did not significantly differ between the IP-C and VTC-H groups.	Multiple More patients completed treatment in the VTC-H program than IP-C (72.9% vs 62.3%, chi sq 5.34, $p < 0.05$. Discharge due to nonattendance and withdrawal from treatment due to dissatisfaction with the program were low in both groups, and differences between groups were not statistically significant.	Mean number of days missed while in treatment & mean number of days attending the program Mean days missed: 1.4 VTC-H vs 1.5 IP-C ($t = 0.86$, NS) Mean days attending: 13.5 VTC-H vs 8.5 IP-C ($t = 7.61$, $p < 0.001$).	CUPSS Upon treatment completion, more than 90% of the patients in the IP-C and VTC-H groups indicated they were very or extremely satisfied with their treatment (97.4% vs 94.0%, $X^2 = 1.24$, NS).	Adverse event tracking No patients in either group attempted or committed suicide.

Abbreviations. ACF=Assessment of Current Functioning; ACQ=Agoraphobic Cognition Questionnaire; BAI=Beck Anxiety Inventory, BDI-II=Beck Depression Inventory; BSQ=Body Sensation Questionnaire; CALPAS=California Psychotherapy Alliance Scale; CI=confidence interval; CMOTS=Client Motivation for Therapy Scale; CPT=cognitive processing therapy; CUPSS=Clinically Useful Patient Satisfaction Scale; DASS=Depression Anxiety Stress Scales; DCCS=Distance Communication Comfort Scale; EMDR=eye movement desensitization and reprocessing; GAD=generalized anxiety disorder; GAD-7=General Anxiety Disorder-7; HR=hazard ratio; IP-C=clinic-based in-person; IP-H=home-based in-person; mPSS=modified PTSD Symptom Scale; NR=not reported; OCD=obsessive-compulsive disorder; OCI-R=Obsessive Compulsive Inventory – Revised; OTT=over the telephone; PAS=Panic and Agoraphobia Scale; PCL-5=PTSD Checklist for DSM-5, PCL-M=PTSD Checklist-Military Version; PCL-S=PTSD Checklist-Specific Version; PDSS-SR=Panic Disorder Severity Scale, Self-Report; PE=prolonged exposure; PHP=partial hospital program; PHQ-9=Patient Health Questionnaire-9; PSWQ=Penn State Worry Questionnaire; PT=post-treatment; PTSD=post-traumatic stress disorder; QIDS=Quick Inventory of Depressive Symptoms; QLESQ=Quality of Life Enjoyment and Satisfaction Questionnaire; RDQ-M=Modified Remission from Depression Questionnaire; SAD=social anxiety disorder; SDS=Sheehan Disability Scale; SEQ=Session Evaluation Questionnaire; SPIN=Social Phobia Inventory; STAI=State-Trait Anxiety Inventory; SUD=substance use disorder; TH=telehealth; VTC-C=clinic-based videoteleconference; VTC-H=home-based videoteleconference; VT-Q=Videoconference Therapy Questionnaire; VTS=Videoconferencing Telepresence Scale; WAI=Working Alliance Inventory; WSAS=Work and Social Adjustment Scale; Y-BOCS-SR=Yale-Brown Obsessive-Compulsive Scale – Self-Report.

QUALITY ASSESSMENT OF INCLUDED PRIMARY STUDIES

ROB 2

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Acierno 2016[1] & Strachan 2012[2]	Low Randomization sequence was generated by a biostatistician using a permuted block randomization scheme with block size varied. Groups did not differ on any demographic variable.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear Treatment fidelity was at or above 90%. 69% of participants completed at least 5 of 8 treatment sessions and had follow-up data; attrition was similar across groups. Analyses conducted with per protocol sample. Unclear co-interventions; included participants receiving other mental health treatment.	High No follow-up data was available for 24% of randomized participants. ITT sample included only participants for whom at least 1 posttreatment or follow-up data point was available. Unclear level or handling of missing data for the analyzed sample.	Unclear Outcome assessors blinded, but outcomes were self-report measures and participants were not blind to treatment condition.	Unclear Design paper specifies superiority analyses conducted with the ITT sample that is not reported in the results.	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Acierno 2017[3] & Gros 2018[4] & Gros 2018[5] & Reich 2021[6] & Yuen 2015[7]	Low Randomization sequence was generated by a biostatistician using a permuted block randomization scheme with block size varied. Groups did not differ significantly on any demographic characteristics or self-report measures of psychopathology.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear Treatment fidelity was at or above 90%. 65% of participants completed at least 6 sessions; attrition was similar across groups. Primary analysis conducted with per protocol sample. Unclear co-interventions; included participants receiving other mental health treatment.	High For participants included in analyses, missing data were estimated using multiple imputation. However, main analysis was conducted with per protocol sample and the ITT sample did not include all participants randomized.	Unclear Outcome assessors blinded, except for outcomes assessed during treatment sessions. Outcomes were self-report measures and participants were not blind to treatment condition.	Unclear Design paper specifies noninferiority analyses will be conducted with both ITT and per protocol samples but results only report per protocol sample.	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Acierno 2021[8] & Gilmore 2020[9] & Lopez 2021[10]	Low 1:1 randomization using REDCap was done by the project coordinator and assignments were saved as they were made and reviewed by the study statistician. There were no significant baseline differences between treatment conditions.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear Treatment fidelity was rated as high. Dose of therapy received did not differ significantly between treatment conditions. Unclear co-interventions.	Low Conducted both ITT and per protocol analyses and ITT analyses include all participants randomized. Appropriate methods were used for imputation of missing data.	Unclear Outcome assessors blinded, but outcomes were self-report measures and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Low

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Alegria 2014[27] & Aguilera 2018[28] & Alcantara 2016[29] & Kafali 2014[30]	Unclear Only states that participants were randomized. Allocation concealment not described. Groups were comparable on sociodemograph-ic characteristics.	Low Cannot blind participants or providers. No deviations from assignment reported.	Low Participants in IP-C group were twice as likely to not initiate care after randomization than participants in telephone group, but ITT analyses included all randomized participants. Procedures were in place to ensure treatment fidelity and standardize delivery of interventions. Participants could not be currently receiving specialty mental health care. Conducted sensitivity analysis on medication use.	Low Of 257 patients randomized, 56 did not complete the 1st follow-up, and 24 more did not complete the 2nd follow-up interview. All participants randomized were included in analyses. Appropriate methods were used for imputation of missing data.	Unclear Outcome assessors blinded, but outcomes were self-report measures and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Some concerns

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Cherestal 2019[75]	Unclear An online randomization website was used to generate a randomization list of 20 subjects. Allocation concealment not described. Groups were comparable with respect to gender, age, and baseline flight anxiety score.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear 4 participants of out 22 randomized did not complete treatment (2 in each group). All other participants completed treatment. Treatment fidelity was not assessed. Participants were required to avoid or cease use of benzodiazepines during exposure sessions; no additional information provided on co-interventions.	High 4 withdrawals were excluded from analyses. Unclear level or handling of missing data for the analyzed sample.	Unclear Outcomes were assessed via self-report measures completed by participants who were not blind to treatment condition.	Unclear Protocol not identified	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Choi 2014[76]	Unclear Only states that participants were randomly assigned. Allocation concealment not described. No significant differences in baseline characteristics between groups.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear 12% of randomized participants dropped out before completing 6 sessions. Procedures included ongoing clinical supervision and fidelity monitoring and there were no differences between groups. Excluded individuals currently involved in psychotherapy; included individuals on antidepressants. Did not assess antidepressant use.	Unclear In addition to the 14 dropouts, 11 participants who completed 6 sessions did not complete the 12- and/or 24-week follow-up. States that intent-to-treat analysis was conducted by n's for analysis are not provided and no participant flow diagram is given.	Unclear Outcomes appear to include self-report measures as well as structured interviews. No information on outcome assessors is provided.	Low Appear to have reported all prespecified findings	Some concerns

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
De Las Cuevas 2006[50]	Unclear Only states that participants were randomly assigned. Allocation concealment not described. No significant differences in baseline characteristics between groups.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear Of 140 participants randomized, 4 dropped out from the VTC-C group and 6 from the IP-C group. Treatment fidelity was not assessed. Unclear co-interventions.	Unclear Number of participants who completed follow-up assessments is not provided. Appears that analyses were conducted with treatment completers, but attrition was low. Unclear level or handling of missing data for the analyzed sample.	Unclear Measures were self-report or clinician-rated with neither participants nor providers blind to treatment condition.	Unclear Protocol not identified	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Egede 2015[32] & Egede 2016[33] & Egede 2017[34] & Egede 2018[35]	Low The randomization sequence was computer-generated using a permuted-block randomization scheme stratified by race with block size varied. Patient allocation was provided in individual sealed envelopes to the study coordinator.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear 20% of participants randomized did not complete all treatment sessions; attrition was similar across groups. Procedures included clinical supervision and fidelity monitoring; all therapists achieved at greater than 90% protocol-specified fidelity. Permitted stabilized use of antidepressants but did not collect data on use.	Low 8.3%, 10%, 12.5%, and 16.7% of data were missing and imputed at week 4, week 8, month 3, and month 12. Did not find any systematic differences in missing and not missing values between groups. Both intention-to-treat and per protocol analyses were conducted.	Unclear Outcome assessors blinded, but outcomes were self-report measures and participants were not blind to treatment condition.	Unclear Appears some secondary outcome domains specified in design paper are not reported.	Some concerns

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Franklin 2017[11]	Unclear Participants randomly assigned using a permuted block procedure. Does not specify block size. Allocation concealment not described.	Low Cannot blind participants or providers. No deviations from assignment reported.	High Out of 10 randomized to VTC-H, 7 did not complete treatment. Out of 7 randomized to VTC-H, 3 did not complete treatment. Differences in dropout rates between groups were statistically significant. Technical problems occurred in the VTC-H group that interfered with the intervention. Treatment fidelity was assessed as high. Unclear co-interventions.	High 6/17 patients completed assessments. Change in outcomes over time across treatment conditions are reported for completers only. Unclear level or handling of missing data for the analyzed sample.	Unclear Outcome assessors blinded, but outcomes were self-report measures and participants were not blind to treatment condition.	Unclear Protocol not identified	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Frueh 2007[12] & Frueh 2007[13]	Unclear Urn randomization was used. Allocation concealment not described. No significant differences in baseline characteristics between groups.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear 8/17 participants in the VTC-C group did not complete treatment; 9/21 participants in the IP-C group did not complete treatment. Rate of attrition was not significantly different between groups. Treatment adherence was assessed and was not statistically significant between groups. Participants were not excluded based on medication use; unclear co-interventions.	High Only 25 of 38 randomized had at least 1 post-baseline assessment; these 25 participants comprise the primary analysis set (completer analysis). Analyses were repeated using a data set for which missing end-of-treatment score for dropouts after the 7-week visit were imputed using the 7-week score.	Unclear Measures were self-report or clinician-rated with neither participants nor providers blind to treatment condition.	Unclear Protocol not identified	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Glassman 2019[14]	Unclear Sequence generation was described in 1 of 2 parent studies. Allocation concealment was not described in either parent study. No significant differences in baseline variables reported in parent studies.	Unclear Cannot blind participants or providers. In both parent studies some randomized individuals did not receive treatment and these individuals were not included in analyses.	Unclear About a quarter of participants in parent studies did not complete treatment. Fidelity was rated high in both parent studies. Both parent studies required stable medication regiments but provided no additional description of co-interventions.	Low Appears that all enrolled participants from each parent study were included in analyses. Appropriate methods handling of missing data.	Unclear Outcome assessors blinded, but outcomes were self-report measures and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Some concerns
Haghnia 2019[15]	Unclear Simple randomization was used. States that participants were matched for age and severity of illness. Allocation concealment not described. Appears to be a statistically significant difference between groups on literacy.	Unclear Cannot blind participants or providers. Six randomized participants did not receive the intervention (2 allocated to IP-C treatment and 4 allocated to VTC-H treatment); no further detail is provided.	Unclear 1 participant in the IP-C group and 2 participants in the VTC-H group discontinued the intervention. Treatment fidelity was not assessed. Participants with comorbid psychiatric diagnoses were allowed and participants continued to receive standard care (no further detail on co-interventions is provided).	High 60/71 participants randomized were included in analyses (did not include dropouts or participants lost to follow-up). Unclear level or handling of missing data for the analyzed sample.	Unclear A questionnaire was provided to participants, who were not blind to treatment condition. Research assistants managing the data collection were blinded.	Low Appear to have reported all prespecified findings	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Hernandez-Tejada 2014[16]	Low Sequence generation and allocation concealment were adequate in both parent studies and there were no significant baseline differences reported in overall sample.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear Treatment fidelity of parent studies was rated as high. Dose of therapy received did not differ significantly between treatment conditions. Unclear co-interventions.	High 68% of treatment dropouts agreed to provide data for either or both outcome measures.	Unclear Outcomes were assessed via self-report measures completed by participants who were not blind to treatment condition.	Low Appear to have reported all prespecified findings	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Hungerbuehler 2016[36]	Unclear States that a randomization list was prepared in Excel but does not specify how sequence was generated. Allocation concealment not described. No significant differences in baseline demographic characteristics between groups, but the VTC-H group showed significantly higher levels of depression and worse mental health status at baseline.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear At 6 months the difference in dropouts between groups was statistically significant. (3/53 participants in VTC-H group and 10/54 in IP-C group discontinued the intervention). Treatment fidelity was not assessed. Unclear co-interventions.	High 8/53 participants in VTC-H group and 14/54 in IP-C group were lost to follow-up and not included in 12-month analyses. Unclear level or handling of missing data for the analyzed sample.	Unclear Outcome measures were self-reported, both during an in-person follow-up consultation with the psychiatrist and afterwards with a web-based questionnaire; participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Liu 2020[17]	Unclear Randomization was stratified by study therapist; no further detail is provided. Allocation concealment not described. There were some baseline differences between treatment groups, which were included as potential covariates in the multivariable analysis.	Low Cannot blind participants or providers. 1 participant randomized to VTC-C and 3 participants randomized to IP-C did not receive the allocated intervention (lost contact with participant before therapy). However, all participants randomized are included in ITT analysis.	Unclear 78/103 in VTC-C group and 68/104 in IP-C group completed at least 9 treatment sessions; dropout rates were not significantly different between groups. Study therapists received weekly consultation but there is no mention of assessment of treatment fidelity. Excluded individuals concurrently receiving psychotherapies targeting PTSD or depression.	Unclear Completer analysis included 65 participants in VTC-C group and 73 in IP-C group who finished baseline and post-treatment visits as well as 9 treatment sessions. Intention-to-treat analyses included all participants randomized. Do not describe handling of missing data.	Low Clinician-rated measure (primary outcome) was administered by independent, blind clinical evaluators. Additional outcomes were self-reported, and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Some concerns

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Lovell 2006[48]	Low Used permuted blocks to randomize patients; did not vary block size. Study therapists contacted the principal investigator to obtain treatment allocation. Baseline symptom severity was comparable between groups.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear 3/36 in IP-C group and 1/36 in telephone group did not complete the intervention. Treatment fidelity was not assessed but therapists had ongoing supervision. Excluded individuals taking antidepressants/anxiolytics for less than 3 months; no additional information on co-interventions.	Unclear 6/36 in IP-C group and 1/36 in telephone group did not complete the 6-month follow-up. States that data were analyzed on an ITT basis but it is unclear whether all randomized individuals were included in analyses. Did not use LOCF to impute data.	Unclear Outcome assessors blinded, but outcomes were self-report measures and participants were not blind to treatment condition. Some participants revealed treatment status to outcome assessor.	Low Appear to have reported all prespecified findings.	Low

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Luxton 2016[37] & Bounthavong 2018[38] & Pruitt 2019[39] & Smolenski 2017[40]	Unclear Randomized in a 1:1 ratio in Stata using permuted blocks, stratified by depression diagnosis and study site; did not vary block size. Allocation concealment not described.	Low Cannot blind participants or providers. 3 participants in each group did not start treatment; ITT analysis includes these individuals.	Low 63% of VTC-H group and 71% of IP-C group completed treatment; difference in dropout between groups was not statistically significant. Treatment fidelity was assessed, and adherence was 98%. Required individuals taking psychoactive medications to be on a stable regimen for at least 30 days. Excluded individuals currently undergoing psychotherapy for depression.	Low Analyses were conducted with both ITT and treatment completer samples. Used regression models that assume data to be missing at random.	Unclear Outcome assessors blinded, but outcomes were self-report measures and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Low

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Maieritsch 2016[18]	Unclear Participants randomized 1:1 using permuted blocks, stratified by gender; did not vary block size. Allocation concealment not described. No significant differences in baseline characteristics between groups.	Unclear Cannot blind participants or providers. 5 individuals randomized to each group did not receive the allocated intervention.	Unclear 38% in the VTC-C group and 43% in the IP-C group who received the allocated intervention dropped out. Treatment fidelity was assessed, and adherence was 96%. Participants on psychotropic medications required to be on a stable dose and maintain that regime during the study; no additional information on co-interventions.	High Study attrition was high (43%). Methods section states that the ITT population was used for analyses, but participant flow diagram indicates the number analyzed was only individuals who completed treatment.	Unclear Does not state whether outcome assessors were blinded, but outcomes were self-report measures and participants were not blind to treatment condition.	Unclear Protocol not identified	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Mohr 2012[41] & Kalapatapu 2014[42] & Stiles-Shields 2014[43] & Stiles-Shields 2015[44]	Low An independent statistician used computer-generated randomization with a 1:1 ratio, stratified by antidepressant status and therapist, with a block size of 4 within each stratum. Randomization was conducted after entry criteria were confirmed. No significant differences in baseline characteristics between groups.	Low Cannot blind participants or providers. All participants received treatment as randomized.	Unclear Significantly fewer participants discontinued treatment in the telephone group (21%) than in the IP-C group (33%). Treatment fidelity was assessed. Excluded individuals who were receiving or planning to receive individual psychotherapy or who had initiated antidepressant pharmacotherapy in the past 10 days.	Low ITT analysis included all randomized individuals. Used appropriate methods for imputation of missing depression scores.	Unclear Outcome assessors blinded, but depression outcomes were self-report measures and participants were not blind to treatment condition.	Unclear After the trial began, changes were made to the planned analyses based on recommendations from the data and safety monitoring board.	Low

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Morland 2014[19] & Morland 2011[20]	Unclear Randomly assigned participants using block randomization, stratified by war era, by an independent off-site clinician. Allocation concealment not described. No significant differences in baseline demographic characteristics between groups.	Unclear Cannot blind participants or providers. 7 individuals in in-person group and 12 in VTC group dropped out between randomization and session 1; these individuals were not included in analyses.	Unclear 78% of participants in the IP-C group and 75% in the VTC-C group attended at least 10 of 12 group sessions. Treatment fidelity was high (99%). Individuals taking psychotropic mediations were required to be on a stable regimen for a minimum of 45 days; no additional information on co-interventions.	Low ITT analysis included all enrolled individuals. Used appropriate methods for imputation of missing values.	Low Clinician-rated measure (primary outcome) was administered by blind assessors. Additional outcomes were self-reported, and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Some concerns

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Morland 2015[21]	Unclear No description of sequence generation or allocation concealment. No significant differences were found between groups on any background variables.	Unclear Cannot blind participants or providers. 23 individuals randomized were not enrolled in study (declined participation or pilot cases); these individuals were not included in analyses.	Unclear 21% of enrolled participants in the IP-C group and 25% in the VTC-C group did not complete treatment. Therapist adherence to the treatment protocol was rated as high. Required a stable psychotropic medication regimen for 45 days prior to study entry; no additional information on co-interventions.	High ITT analysis included all enrolled individuals, but not all randomized participants were 'enrolled'. Used appropriate methods for imputation of missing values.	Low Clinician-rated measure (primary outcome) was administered by blind assessors. Additional outcomes were self-reported, and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Morland 2020[22]	Unclear Randomization was balanced within therapists, using blocks of decreasing sizes. Allocation concealment not described. No significant differences were found between groups on any background variables except for the IP-H group having more Navy Veterans.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear 79% of IP-H group, 62% of VTC-H, & 46% of VTC-C completed treatment. Therapist adherence to the treatment protocols was rated as 95%. Required a stable psychotropic medication regimen for 60 days prior to study entry. Excluded participants undergoing concurrent PTSD or exposure treatment.	Low ITT analysis included all enrolled individuals. Used appropriate methods for imputation of missing values.	Low Clinician-rated measure (primary outcome) was administered by blind assessors. Additional outcomes were self-reported, and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Low

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Morland 2022[23]	Unclear Randomization was done by an independent, off-site statistician; no further detail provided on sequence generation or allocation concealment. Some baseline differences between groups (average relationship length).	Low Cannot blind participants or providers. Five couples received no treatment. Two couples randomized to IP-C received a portion of sessions remotely. These participants were included in analyses in the group they were randomized to & sensitivity analysis was conducted excluding these couples.	Unclear 74% of IP-C group and 70% of VTC-H group completed treatment. Treatment fidelity was rated as over 90%. Required a stable psychotropic medication regimen for 2 months prior to study. Participants could not receive other psychotherapy for PTSD or conjoint psychotherapy during the study. No additional information on co-interventions.	Low Rates of participants missing data on 1 or more outcome variables were 4% at baseline and 58% at the 6-month follow-up. ITT analysis included all enrolled individuals. Used appropriate methods for imputation of missing values.	Low Clinician-rated measure (primary outcome) was administered by blind assessors. Additional outcomes were self-reported, and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Low

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
O'Reilly 2007[51]	Unclear First group of participants was randomly assigned by coin flip; used block randomization for subsequent participants. Allocation concealment not described. No significant baseline differences between groups.	Unclear Cannot blind participants or providers. 3% of participants in IP-C group and 7% of participants in VTC-H group did not receive treatment. These participants were not included in analyses.	Unclear Treatment dropout and fidelity are not reported. No information on co-interventions provided.	High Out of the participants who received treatment, 40% in IP-C group and 38% in VTC-H group had incomplete follow-up assessments. Analyses on psychiatric symptoms only include participants with complete follow-up assessments.	Unclear Outcomes were self-report measures and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Peterson 2022[24]	Unclear Equipoise stratified randomization where participants could opt out of 1 delivery modality. No additional information is provided on sequence generation or allocation concealment. No significant baseline differences between groups except for years of service.	Low Cannot blind participants or providers. 8 participants were randomized and never began treatment; these participants were included in ITT analyses.	Unclear 43% in IP-C group, 25% in IP-H group, & 34% in VTC-H group dropped out of treatment. Treatment fidelity was assessed, and adherence was rated as 99%. Required stable medication regimen; no additional information on co-interventions.	Unclear 39% in IP-C group, 28% in IP-H group, & 34% in VTC-H group did not complete at least 1 follow-up. ITT analyses were used and included all participants randomized. Unclear handling of missing data.	Low Clinician-rated measure was administered by blind assessors. Additional outcomes were self-reported, and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Some concerns

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Ruskin 2004[45] & Rhein 2001[46]	Unclear Variable block randomization; stratified based on age and depression severity. Allocation concealment not described. No significant baseline differences between groups.	Low Cannot blind participants or providers.10 post-randomization exclusions of ineligible participants, 2 eligible but did not start treatment.	Unclear 27% of VTC-C group and 30% of IP-C group dropped out of the study. Treatment fidelity was not assessed. Excluded individuals who had been receiving pharmacological treatment for depression for more than a month prior; no additional information on co-interventions.	Unclear Extent of missing data unclear; no participant flow diagram. Patients lost to follow-up were retained in the analysis by using their last observable score for all remaining time points.	Unclear Outcomes were self-report measures and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	Some concerns
Stubbings 2013[52]	Unclear Simple random allocation achieved by generating a randomized number list. Allocation concealment not described. No significant baseline differences between groups.	Low Cannot blind participants or providers. No deviations from assignment reported.	Low 81% of participants completed the full course of treatment. Fidelity was monitored and treating provider underwent weekly supervision. Excluded individuals currently receiving psychotherapy and a record of medication usage was taken.	High 65% of participants completed the follow-up data. Final analysis included only participants who completed 8 or more sessions.	Unclear Outcomes were self-report measures and participants were not blind to treatment condition.	Unclear Protocol not identified	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
Watts 2020[49]	Low Randomization sequence generated with an iPhone random number generator. The assignment list was concealed to all researchers except research coordinators. No significant baseline differences between groups.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear 14% in IP-C and 23% in VTC-C dropped out. Treatment fidelity was assessed, and adherence was rated as over 90%. Participants could not receive any other psychotherapy or start/stop antidepressant medication; no additional information on co-interventions.	High 33 randomized participants (22%) were not included in analyses; includes participants lost to attrition, participants for whom GAD was no longer the primary diagnosis, and participants with more than 20% missing data on primary outcome measure. Missing data were replaced using the mean from the previous and following information.	Unclear Outcomes were self-report measures and participants were not blind to treatment condition.	Unclear Protocol not identified	High

Author Year	Risk of Bias from Randomization Process	Risk of Bias from Deviation from Intended Interventions (Assignment)	Risk of Bias from Deviation from Intended Interventions (Adherence)	Risk of Bias from Missing Outcome Data	Risk of Bias in Measurement of Outcome	Risk of Bias in Selection of Reported Result	Overall Bias (High, Low, Unclear)
White 2021[25]	Low Block randomization was conducted at a centralized location. There were no significant baseline differences between groups.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear Attrition was high but similar for both treatment groups. Treatment fidelity was rated high. Both parent studies required stable medication regimens but provided no additional description of co-interventions.	High Analysis only included participants with full outcome data (49% of all randomized to both studies) on measure of interest (satisfaction).	Unclear Outcome assessors blinded, but outcomes were self-report measures and participants were not blind to treatment condition.	Low Appear to have reported all prespecified findings	High
Ziemba 2014[26]	Unclear Individuals randomized 1:1 through a computer algorithm. Allocation concealment not described. Baseline differences between groups not reported.	Low Cannot blind participants or providers. No deviations from assignment reported.	Unclear Dropout alone is not reported but 72% of participants completed treatment and all assessments. Treatment fidelity was assessed and was rated as high. Required stable medication regimen; no additional information on co-interventions.	High 28% of participants did not complete treatment and all assessments and were not included in analyses.	Low Clinician-rated measure was administered by blind assessors. Additional outcomes were self-reported, and participants were not blind to treatment condition.	Unclear Protocol not identified	High

Abbreviations. GAD=generalized anxiety disorder; IP-C=clinic-based in-person; IP-H=home-based in-person; ITT=intent-to-treat; LOCF=last observation carried forward; PE=prolonged exposure; PTSD=post-traumatic stress disorder; RCT=randomized controlled trial; VTC-C=clinic-based videoteleconference; VTC-H=home-based videoteleconference.

ROBINS-I

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Bean 2022[66]	Serious There were no significant differences between groups on demographic or clinical variables at baseline, but baseline symptom severity was not compared and there was no adjustment for confounders in analyses.	Serious Selected only participants who completed the treatment program during a specified period and provided complete symptom assessment data. All participants were followed from the start to the end of the intervention.	Moderate Differential misclassification of intervention status is unlikely. However, participants were assigned based on the date they completed the program (not recorded at the start of the intervention).	No information Medication use and prior psychological treatment were not significantly different between groups at baseline. Since groups were defined by completion date, participants in the group who completed the program after the COVID-19 transition date may have started treatment prior to that date and received some portion of treatment in person (crossover) - data on this are not provided.	Serious Participants were excluded due to missing data and a significantly greater percentage of participants in the IP-C group provided complete data than patients in the VTC-H group. There is no evidence that results were robust to the presence of missing data.	Low Outcomes were self-reported, but measurement was comparable across groups. In an RCT participants could not be blinded to intervention group.	Low Appears unlikely that additional, unreported measure-ments/ analyses were done	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Bouchard 2004[62]	Serious Analyses did not adjust for confounding factors, in particular those associated with rural vs urban status (or other site differences). The 2 groups differed significantly on mean number of years in school.	Low Participants were not selected based on participant characteris-tics observed after the intervention, and participants are followed from the start of the intervention.	Low Participants were assigned based on the region (remote vs local site) they were referred from.	No information No deviations from assignment reported. Excluded individuals receiving other psychological treatment and required antidepressant and benzodiazepine users to be on a stable regimen. Unknown whether medication use was balanced across groups. Appears that the intervention was implemented as intended.	No information No information is provided on missing data. It is unclear whether the n's provided are the number of participants randomized and the number of participants included in analyses is not specified.	Low Outcomes were self-reported, but measurement was comparable across groups. In an RCT participants could not be blinded to intervention group.	Moderate Alliance measures were added to the protocol after the study began which meant some were missing for the IP-C group - as a result none of the alliance results are reported for the IP-C condition.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Bouchard 2020[63]	Moderate There were no significant differences between groups on certain demographic or clinical variables at baseline, but differences in baseline symptom severity and medication use were not reported. Analyses were conducted controlling for gender and presence of a comorbid disorder only. Do not control for other potential confounders.	Low Participants were not selected based on participant characteris-tics observed after the intervention, and participants are followed from the start of the intervention.	Low Participants were assigned based on the region (remote vs local site) they were referred from.	No information Flowchart indicates that a portion of randomized participants in each group did not receive the allocated intervention. Do not report whether there were significant differences between completers and dropouts. Treatment fidelity was assessed. Excluded individuals currently receiving a psychological treatment and required antidepressant and benzodiazepine users to be on a stable regimen.	Low Analyses included all participants randomized (ITT) and online supplement includes results with imputed values for missing data (these analyses did not affect findings).	Low Outcomes were self-reported, but measurement was comparable across groups. In an RCT participants could not be blinded to intervention group.	No information Study protocol was registered retrospect-tively and does not specify all outcomes.	Moderate

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
				Unknown whether medication use was balanced across groups.				

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Gannon 2021[69]	Moderate There were no significant differences between groups on sociodemographic and clinical characteristics assessed, but initial anxiety scores during the COVID-19 period (telehealth) were significantly higher than the pre-COVID-19 (in-person) period. Examined the effect of age, gender, race, diagnosis, IOP type, insurance type, zip code, # assessments completed. Do not account for confounders within a statistical model.	Moderate The 2 groups were selected during different time periods and the beginning and end of the intervention do not coincide.	Low Differential misclassification of intervention status is unlikely. Participants were assigned based on the date they started the program.	No information Other than # weeks of patient enrollment, no information on dropout is provided. No information on co-interventions is provided.	Moderate In instances where only 1 assessment was available, data was imputed using the "last observation carried forward." Do not report how many participants had data for both assessments.	Serious Outcomes were self-reported. Scores were taken from the electronic medical record; assessment methods differed between groups, with the therapists facilitating office-based assessments in the IP-C group and participants completing assessments on their own outside of groups.	Low Appears unlikely that additional, unreported measurements/ analyses were done	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Germain 2009[53] & Germain 2010[54]	Moderate There were no significant differences between groups on sociodemograph-ic variables and severity and frequency of posttraumatic symptoms. Analyses do not control for confounders.	Low Participants were not selected based on participant characteris-tics observed after the intervention, and participants are followed from the start of the intervention (although not all participants had the waiting period before treatment).	No information Remote participants were assigned to the VTC-C condition, but it is unclear how participants recruited in Montreal were assigned to treatment condition.	Serious Some randomized participants were excluded from the therapy. Very high attrition. Reports that dropouts and treatment completers were not different on sociodemographi c and baseline symptom variables but does not report whether dropout was significantly different between the groups. Treatment fidelity was assessed (95%). Required that participants agree to not participate in other forms of psychotherapy during treatment.	Serious Excluded data on 25% of the initial sample due to dropout. No information on handling of missing data and n's are not provided for analyses.	No information Does not specify whether students who administered the clinical interview were blind to treatment condition. Additional outcomes were self-reported, but measurement was comparable across groups. In an RCT participants could not be blinded to intervention group.	No information Non-randomized trial; no trial protocol identified.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
				Required users of psychotropic medications to have stable regimens. Unknown whether medication use was balanced across groups.				
Gros 2011[55]	Moderate The comparison group was matched on demographic characteristics but not on symptom severity and only included treatment completers. Analyses demonstrated no differences in baseline symptoms between groups. When comparing groups only used treatment completers.	Serious Selection of comparison group was based on completion of treatment. It is unclear whether participants in the VTC-C group were recruited consecutive-ly.	Low Differential misclassification of intervention status is unlikely. Treatment group depended on where patients receive care.	Moderate Treatment provided was routine clinical practice. There was a significant difference between groups in number of sessions completed.	Serious Main analyses of symptom change were conducted only with treatment completers and attrition was high. Extent and handling of missing data is unclear.	Low Outcomes were self-reported, but measurement was comparable across groups. In an RCT participants could not be blinded to intervention group.	Low Outcome measures for the study were the measures assessed as part of routine care that were shared between both settings.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Hammond 2012[70]	Moderate Analyses included an adjusted model controlling for baseline symptom severity, provider site, and treatment duration. Also included analyses that used sampling methods based on propensity scores (covariates included demographics, symptom severity, work, social adjustment, service level predictors, presence of phobias, employment, benefit status, receipt of psychotropic	Low Appear to have selected from sample of all individuals referred to the program at eligible sites during the specified period who met the criteria.	Low Differential misclassification of intervention status is unlikely. Treatment group depended on treatment modality and individuals who received care via both modalities were excluded from analyses.	Low Excluded individuals who attended less than 2 sessions. Excluded individuals scheduled to receive high-intensity interventions. Intervention was delivered as routine clinical practice.	Serious Excluded individuals with less than 2 sessions where assessments were completed. Only 68% of individuals allocated to the low-intensity intervention who attended at least 2 sessions had information on baseline and endpoint measures and baseline covariates and were included in the study. Within this group, only individuals who had completed	Low Outcomes were self-reported as part of routine clinical practice. Patients were not aware that the study would take place. In an RCT participants could not be blinded to intervention group.	Low Outcome measures for the study were the measures assessed as part of routine care that were shared between both settings.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
	medication, referral source, study site).				treatment by the time of data extraction were included in analyses. Missing baseline data were imputed using correspond-ding variables at the second visit.			

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Khatri 2014[71]	Serious Participants in each group did not differ on the initial clinical interview (DSM-IV). Does not report if differences between groups on demographic variables were significant. Potential confounders were not addressed in analyses.	Low Included individuals meeting study criteria out of those who were referred to the study by their physicians.	Low Participants chose the intervention format they felt more comfortable using.	No information Three participants withdrew from the study. Adherence was rated using a coding system and was high. Participants agreed to remain on stable doses of any prescribed psychotropic medication; no additional information on co-interventions is provided and data on medication use were not reported.	Moderate Three withdrawals were not included in analysis of within-person pre-post-depression symptom change. No description of extent or handling of missing data.	Low Outcomes were self-reported, but measurement was comparable across groups. In an RCT participants could not be blinded to intervention group.	No Information Non-randomized pilot trial; no trial protocol identified.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Knowlton 2021[56]	Moderate There were no significant differences between groups on reported demographic and clinical variables. Data on several process variables were collected to serve as possible covariates.	Low Appear to include all individuals who met criteria during the given period within the catchment area.	Low Participants received services via a given modality based on geographic location and preference.	Serious 53% of participants who initiated treatment did not complete a minimum of 8 sessions; there were no significant differences in treatment completion between groups. Excluded participants with prior engagement in PTSD treatment; no information on co-interventions is provided.	Serious Excluded participants who did not complete at least 8 sessions (minimum treatment dosage). Use of outcome tracking measures during treatment sessions was inconsistent. No description of extent or handling of missing data.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Liou 2022[72]	Moderate Differences in change scores between groups were estimated using multiple linear regression, adjusting for factors found to significantly differ between groups.	Moderate Retrospective chart review of all adult patients seen by 1 provider at a primary care center during a specified period. Groups differed on timing of the beginning and end of the intervention.	Low It appears that patients received virtual treatment because of the COVID-19 pandemic.	No information Patients with 1 or more EMDR session were classified as receiving EMDR; patients with 1 or more virtual EMDR visit were classified as receiving virtual EMDR. However, only 1 patient (0.3%) received both virtual and in-person EMDR, and most patients (73.5%) had 3-5 EMDR sessions. No information on co-interventions.	No information No information on missing data is provided.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Low Outcome measures for the study were the measures assessed as part of routine care that were shared between settings.	No information

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
LoSavio 2021[57]	Moderate Employed propensity matched scoring using demographic differences (age, education, race/ethnicity, relationship status, trauma type, comorbid substance use) and baseline PCL-5 score.	Moderate Appears to include all patients that were treated by clinicians participating in the treatment training and implementation program. Likely that timeframe of intervention differed between groups.	Low Treatment modality was determined by patient/provider preference or as needed in response to the COVID-19 pandemic.	Moderate Overall dropout was 25%; patients receiving care via telehealth were significantly more likely to complete treatment than patients receiving care in-person. Treatment was delivered as part of routine clinical care. No information on co-interventions.	Low Growth curve models include all ITT data regardless of how many sessions patients completed.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Low Outcome measures for the study were the measures assessed as part of routine care that were shared between settings.	Moderate

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Milosevic 2022[64]	Serious There were significant differences between groups on some demographic and clinical variables. Conducted supplementary analyses controlling for comorbidities and baseline depression scores.	No information Unclear whether all eligible patients were selected for the study.	Low Treatment modality was determined by date (before or after COVID-19 pandemic)	No Information Dropout was similar between groups. Treatment was delivered as part of routine clinical care. No information on co-interventions.	No information States that an ITT approach was taken for symptom outcome measures. Hierarchical linear modeling was used but the extent and handling of missing data is not described.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Low Outcome measures for the study were the measures assessed as part of routine care that were shared between settings.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Pinciotti 2022[65]	Moderate Patients receiving telehealth treatment were more likely to have GAD and mood disorders compared to those receiving pre-pandemic in-person treatment. Demographic variables and comorbidities were extracted for use as potential covariates. Other covariates included level of care, number of diagnoses, length of stay. Analyses were conducted with participants matched for admission scores.	Low Appears that patients who met criteria and were treated in the specified period were included.	Low Treatment modality was determined by date (before or after COVID-19 pandemic).	No information Treatment was delivered as part of routine clinical care, but telehealth treatment had to be modified in some ways due to social distancing requirements. No information on co-interventions.	Moderate In matching, the IP-C group was trimmed to meet the sample size of the VTC-H group. Participants with missing progress data were not included in linear mixed model.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Low Outcome measures for the study were the measures assessed as part of routine care that were shared between settings.	Moderate

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Ritchie 2007[61]	Critical There were some important differences noted between groups (*eg*, the VTC-C group was all active-duty soldiers who walked to the therapy location and the IP-C group were all spouses of military members who had to drive) that were not controlled for in analyses.	Low Participants were referred to the study by health care providers or were self-referred and individuals meeting study criteria were included.	Low Patients were recruited for the VTC-C group or IP-C group based on their preference for location/time and not preference for delivery method.	No information Treatment compliance was 98% for the VTC-C group and 71% for the IP-C group; does not report whether this difference is significant. Group CBT was conducted as the standard of care. Participants could not be currently receiving other psychotherapy treatment. Required users of psychotropic medication to have a stable regimen; rates of medication use are not reported.	No information All participants appear to be included in analyses. No information is provided on the extent or handling of missing data.	Low Outcomes were self-reported but measurement was comparable across groups. In an RCT participants could not be blinded to intervention group.	No information Non-randomized pilot trial; no trial protocol identified.	Critical

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Tuerk 2010[58]	Serious A reference sample of individuals who received IP-C treatment was selected but was not matched. There were no significant differences in baseline symptom severity between groups. Do not report whether differences in demographics are significant. Do not control for the potential confounding effects of rural status.	Low Patients were referred by mental health providers and case managers and individuals who met PTSD criteria were invited to participate in treatment.	Low Veterans in rural areas were given the option to receive therapy via telehealth.	No information The treatment completion rate was 83% for the IP-C group and 75% for the VTC-C group; group differences are not examined. Some elements of exposure therapy (*ie*, exercises where the provider accompanies the patient outside of the treatment setting) were not feasible in the VTC-C group. No information on co-interventions.	Moderate Pre- and post-treatment effects are only examined for patients who completed at least 5 treatment sessions. No information is provided on the extent or handling of missing data.	Low Outcomes were self-reported, but measurement was comparable across groups. In an RCT participants could not be blinded to intervention group.	No Information Non-randomized pilot trial; no trial protocol identified.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Vakkalanka 2022[67]	Moderate Measured and accounted for a range of potential confounders including buprenorphine availability/ access, sociodemograph-ic and geographic factors, comorbidities, medication use, health care utilization, prior telehealth use.	Low Included Veterans diagnosed with OUD and treated with buprenorphin e across all facilities within the VHA between 2008 and 2017.	Low Telehealth and in-person encounters were treated as exposures. Used stop codes indicating clinic-to-home or clinic-to-clinic video telehealth to define telehealth encounters.	Moderate Treatment included prescription of buprenorphine. Telehealth group only needed to have a single telehealth encounter but secondary analysis looks at individuals who only had telehealth encounters. Defined treatment discontinuation as a gap greater than 14 days. 66.5% discontinued treatment within 1 year. Included data on past year psychiatric medication use and past 2-year health care	Moderate Did not include individuals missing geographic information. Appear to have used statistical models that account for missing data.	Low Treatment utilization data obtained from VHA records.	Low Treatment utilization data obtained from VHA records.	Moderate

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
				utilization as covariates.				
Valentine 2020[59]	Serious Demographic characteristics and comorbidities were not significantly associated with treatment modality. For those assessed with the CAPS-IV, there were no significant baseline differences symptom severity, but for those assessed with CAPS-5, patients seen in-person has significantly lower scores than those using CVT.	Low Appear to include all Veterans seen within a PTSD specialty clinic during the specified period.	Low Veterans chose their preferred treatment modality in collaboration with providers.	No information No significant differences between groups in treatment completion rates. Include both home and clinic telehealth. Veterans chose to receive PE or CPT which was provided as routine care. Session notes were reviewed and sessions that did not follow treatment protocols did not count. No information on co-interventions.	Low As the main analysis was on treatment completion, all individuals were included.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Low Outcome measures for the study were the measures assessed as part of routine care that were shared between settings.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Wierville 2016[60]	Moderate Service era, service connection and PCL-S scores at pre-treatment were significantly different between groups. These variables were included in multilevel modeling analyses as covariates.	Moderate Inclusion of in-person patients was based on CAPS, while inclusion of telehealth patients was based on PCL-S and chart notes.	Low Veterans chose their preferred treatment modality.	No information Dropout rates were not significantly different between groups. Only patients who completed intake and at least 1 session were included in the study. Veterans chose PE or CPT as part of routine clinical practice. No information on co-interventions.	Low Missing data was present on 4% of the values in the overall dataset, which included 51% of participants having missing data on at least 1 variable. Multiple imputation was used to account for missing data.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Low Outcome measures for the study were the measures assessed as part of routine care that were shared between settings.	Moderate

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Zheng 2017[68]	Serious No statistical difference was detected between groups for demographic factors. Generalized estimating equations were used to adjust for covariates. Other potential confounders were not addressed (*eg*, rural status).	Low Included patients under care of a psychiatrist who provided care through both settings during a specified period. Included patients with OUD receiving MAT services in weekly or biweekly groups.	Low Treatment group was based on where Veterans lived; the telepsychiatry clinic served 2 rural counties.	Moderate Group therapy was provided face-to-face for both groups; medication management was provided through both telepsychiatry and face-to-face. There were no attempts to ensure the equivalence of group therapy at both sites, although all therapists use a CBT-based therapeutic model.	Serious States there were several instances of missing data due to incomplete records. Do not provide any additional information on extent or handling of missing data.	Low Utilization data obtained via chart review.	Low Data consisted of treatment information regularly recorded for all patients in the program.	Serious

Author Year	Bias Due to Confounding? (Low, Moderate, Serious, Critical, No Information)	Selection Bias (Low, Moderate, Serious, Critical, No Information)	Bias in Classification of Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Departures from Intended Interventions (Low, Moderate, Serious, Critical, No Information)	Bias Due to Missing Data? (Low, Moderate, Serious, Critical, No Information)	Bias Due to Measurement of Outcomes? (Low, Moderate, Serious, Critical, No Information)	Bias in the Selection of Reported Results (Low, Moderate, Serious, Critical, No Information)	Overall Bias (Low, Moderate, Serious, Critical, No Information)
Zimmerman 2017[73]	Serious The IP-C comparison group was not matched; selected every other patient during a specified period. There were significant differences between groups on # widowed and # less than high school graduate. Mean number of current diagnoses was significantly higher in the telehealth group.	Low Included patients who were not previously treated in the program and underwent consent procedures.	Low Treatment modality was determined by date (before or after COVID-19 pandemic)	Moderate There were some differences between the VTC-H and IP-C treatment groups. Significantly more patients completed treatment in the VTC-H program than the IP-C program. No information on co-interventions.	Serious Rate of completion of patient satisfaction measure and treatment outcome measure was significantly lower in the VTC-H cohort. No description of handling of missing data.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Low Outcomes were self-reported as part of routine clinical practice. In an RCT participants could not be blinded to intervention group.	Serious

Abbreviations. CAPS-5=Clinician Administered PTSD Scale; CBT=cognitive behavioral therapy; CPT=cognitive processing therapy; EMDR=eye movement desensitization and reprocessing; GAD=generalized anxiety disorder, IOP=Intensive Outpatient Program; IP-C=clinic-based in-person; ITT=intent-to-treat; MAT=medication assisted treatment; OUD=opioid use disorder; PE=prolonged exposure; PTSD=posttraumatic stress disorder; VTC-C=clinic-based videoteleconferencing; VTC-H=home-based videoteleconferencing.

STRENGTH OF EVIDENCE FOR INCLUDED STUDIES

Strength of Evidence for KQ 1

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
PTSD									
PTSD symptom severity	VTC-H vs IP-C	Individual psychotherapy	Acierno 2016,[1] Acierno 2017,[3] Acierno 2021,[8] Knowlton 2021,[56] Peterson 2022,[24] White 2021[25]	Low to high	Direct	Consistent	Imprecise	Unknown	Low PTSD symptom severity appears similar after VTC-H and IP-C delivery of individual psychotherapy.
		Cognitive-behavioral conjoint therapy	Morland 2022[23]	Low	Direct	Unknown	Unknown	Unknown	Low PTSD symptom severity may be similar after VTC-H and IP-C delivery of couples' psychotherapy.
	VTC-C vs IP-C	Individual psychotherapy	Germain 2009,[53] Gros 2011,[55] Knowlton 2021,[56] Liu 2020,[17] Maieritsch 2016,[18] Morland 2015,[21] Tuerk 2010,[58] Wierwille 2016,[60] Ziemba 2014[26]	High	Direct	Inconsistent	Imprecise	Unknown	Low It is unclear whether PTSD symptom severity differs between VTC-C and IP-C delivery of individual psychotherapy.
		Group psychotherapy	Frueh 2007,[13] Morland 2014[19]	Moderate to high	Direct	Inconsistent	Imprecise	Unknown	Low It is unclear whether PTSD symptom severity differs between VTC-C and IP-C delivery of group psychotherapy at post-treatment. There may be

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
									no difference in PTSD symptom severity at 3-month follow-up.
PTSD diagnosis	VTC-H vs IP-C	Individual psychotherapy	Acierno 2017[3]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether the rate of PTSD remission differs between VTC-H and IP-C delivery of individual psychotherapy.
	VTC-C vs IP-C	Individual psychotherapy	Germain 2009[54]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether the rate of PTSD remission differs between VTC-C and IP-C delivery of individual psychotherapy.
Depression symptom severity	VTC-H vs IP-C	Individual psychotherapy	Acierno 2016,[1] Acierno 2017,[3] Acierno 2021,[8] Knowlton 2021,[56] Peterson 2022[24]	Low to high	Direct	Inconsistent	Imprecise	Unknown	Low It is unclear whether depression symptom severity differs between VTC-H and IP-C delivery of individual psychotherapy immediately post-treatment and at 3-month follow-up. Depression symptom severity may be similar 6–12 months after treatment.
	VTC-C vs IP-C	Individual psychotherapy	Germain 2009,[77] Gros 2011,[55] Knowlton 2021,[56] Liu 2020,[17] Maieritsch 2016,[18] Tuerk 2010,[58] Wierwille	High	Direct	Inconsistent	Imprecise	Unknown	Low It is unclear whether depression symptom severity differs between VTC-C and IP-C delivery of individual psychotherapy.

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
			2016,[60] Ziemba 2014[26]						
		Individual psychotherapy	Frueh 2007[13]	High	Direct	Unknown	Imprecise	Unknown	Insufficient It is unclear whether depression symptom severity differs between VTC-C and IP-C delivery of group psychotherapy.
Anxiety symptom severity	VTC-H vs IP-C	Individual psychotherapy	Acierno 2016,[1] Acierno 2017[3]	High	Direct	Unknown	Imprecise	Unknown	Low Anxiety symptom severity may be similar after VTC-H and IP-C delivery of individual psychotherapy based on preliminary findings.
	VTC-C vs IP-C	Individual psychotherapy	Germain 2009,[53] Ziemba 2014[26]	High	Direct	Consistent	Unknown	Unknown	Low Anxiety symptom severity may be similar after VTC-C and IP-C delivery of individual psychotherapy.
Adverse events	VTC-H vs IP-C	Individual psychotherapy	Peterson 2022[24]	Moderate	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether the occurrence of adverse events differs between VTC-H and IP-C delivery of individual psychotherapy.
	VTC-C vs IP-C	Individual psychotherapy	Tuerk 2010[58]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether the occurrence of adverse events differs between VTC-C and IP-C delivery of individual psychotherapy.
		Group psychotherapy	Morland 2014[19]	Moderate	Direct	Unknown	Unknown	Unknown	Insufficient

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
									It is unclear whether the occurrence of adverse events differs between VTC-C and IP-C delivery of group psychotherapy.
Depression									
Depression symptom severity	VTC-H vs IP-C	Individual psychotherapy	Egede 2015,[32] Luxton 2016[37]	Low to moderate	Direct	Inconsistent	Imprecise	Unknown	Low It is unclear whether depression symptom severity differs between VTC-H and IP-C delivery of individual psychotherapy. One study with low risk of bias found that depression symptom severity decreased less for the VTC-H group than the IP-C group.
		Psychiatry	Hungerbuehler 2016[36]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether depression symptom severity differs between VTC-H and IP-C delivery of psychiatry.
	Telephone vs IP-C	Individual psychotherapy	Alegria 2014,[27] Mohr 2012[41]	Low to moderate	Direct	Inconsistent	Unknown	Unknown	Low Depression symptom severity may be similar after telephone and IP-C delivery of individual psychotherapy immediately post-treatment, but it is unclear whether depression symptom severity differs between telephone and IP-C over time.

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
	VTC-C vs IP-C	Group psychotherapy	Ritchie 2007[61]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether depression symptom severity differs between VTC-C and IP-C delivery of group psychotherapy.
		Psychiatry	Ruskin 2004[45]	Moderate	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether depression symptom severity differs between VTC-C and IP-C delivery of psychiatry.
Anxiety symptom severity	VTC-H vs IP-C	Individual psychotherapy	Luxton 2016[37]	Low	Direct	Unknown	Imprecise	Unknown	Insufficient It is unclear whether anxiety symptom severity differs between VTC-H and IP-C delivery of individual psychotherapy.
	VTC-C vs IP-C	Psychiatry	Ruskin 2004[45]	Moderate	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether anxiety symptom severity differs between VTC-C and IP-C delivery of psychiatry.
PTSD symptom severity	VTC-H vs IP-C	Individual psychotherapy	Luxton 2016[37]	Low	Direct	Unknown	Imprecise	Unknown	Insufficient It is unclear whether PTSD symptom severity differs between VTC-H and IP-C delivery of individual psychotherapy.
Adverse events	VTC-H vs IP-C	Individual psychotherapy	Egede 2015,[32] Luxton 2016[37]	Low to moderate	Direct	Unknown	Unknown	Unknown	Low It is unclear whether the occurrence of adverse events differs between VTC-H and IP-C delivery

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
									of individual psychotherapy.
	Telephone vs IP-C	Individual psychotherapy	Mohr 2012[41]	Low	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether the occurrence of adverse events differs between telephone and IP-C delivery of individual psychotherapy.
Anxiety									
Anxiety symptom severity	VTC-H vs IP-C	Group psychotherapy	Milosevic 2022[64]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether anxiety symptom severity differs between VTC-H and IP-C delivery of group psychotherapy.
		PHP/IOP	Pinciotti 2022[65]	Moderate	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether OCD symptom severity differs between VTC-H and IP-C delivery of PHP/IOP treatment.
	Telephone vs IP-C	Individual psychotherapy for phobia	Cherestal 2019[75]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether anxiety symptom severity differs between telephone and IP-C delivery of individual psychotherapy for phobia.
		Individual psychotherapy for OCD	Lovell 2006[48]	Low	Direct	Unknown	Precise	Unknown	Low OCD symptom severity may be comparable after telephone and IP-C delivery of individual psychotherapy for OCD.

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
	VTC-C vs IP-C	Individual psychotherapy for anxiety	Watts 2020[49]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether anxiety symptom severity differs between VTC-C and IP-C delivery of individual psychotherapy for anxiety.
		Individual psychotherapy for panic disorder with agoraphobia	Bouchard 2004,[62] Bouchard 2020[63]	Moderate to high	Direct	Inconsistent	Unknown	Unknown	Very Low It is unclear whether anxiety-related outcomes differ between VTC-C and IP-C delivery of individual psychotherapy for panic disorder with agoraphobia.
Depression symptom severity	VTC-H vs IP-C	PHP/IOP	Pinciotti 2022[65]	Moderate	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether depression symptom severity differs between VTC-H and IP-C delivery of PHP/IOP treatment.
	Telephone vs IP-C	Individual psychotherapy	Lovell 2006[48]	Low	Direct	Unknown	Precise	Unknown	Low Depression symptom severity may be similar after telephone and IP-C delivery of individual psychotherapy for OCD.
	VTC-C vs IP-C	Individual psychotherapy	Bouchard 2004,[62] Bouchard 2020[63]	Moderate to high	Direct	Consistent	Unknown	Unknown	Low Depression symptom severity may be similar after VTC-C and IP-C delivery of individual psychotherapy for panic disorder with agoraphobia.
Adverse events	VTC-C vs IP-C	Individual psychotherapy	Watts 2020[49]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether the occurrence of adverse

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
									events differs between VTC-C and IP-C delivery of individual psychotherapy for anxiety.
Substance use disorder									
Mental health symptoms	VTC-H vs IP-C	IOP	Bean 2022[66]	High	Direct	Unknown	Precise	Unknown	Insufficient It is unclear whether mental health symptoms differ between VTC-H and IP-C delivery of IOP treatment for co-occurring SUD and mental health diagnosis.
Abstinence	VTC-C vs IP-C	Individual psychotherapy	Zheng 2017[68]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether abstinence differs between VTC-C and IP-C delivery of medication-assisted treatment.
Multiple mental health conditions									
Mental health symptoms	VTC-C vs IP-C	Individual psychotherapy	Stubbings 2013[52]	High	Direct	Unknown	Imprecise	Unknown	Insufficient It is unclear whether mental health symptoms differ between VTC-C and IP-C delivery of individual psychotherapy.
		Psychiatry	De Las Cuevas 2006[50]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether mental health symptoms differ between VTC-C and IP-C delivery of psychiatry.
Depression symptom severity	VTC-H vs IP-C	Individual psychotherapy	Liou 2022[72]	Unclear	Direct	Unknown	Imprecise	Unknown	Insufficient It is unclear whether depression symptom

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
									severity differs between VTC-H and IP-C delivery of individual psychotherapy.
		Group psychotherapy	Khatri 2014[71]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether depression symptom severity differs between VTC-H and IP-C delivery of group psychotherapy.
		IOP	Gannon 2021[69]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether depression symptom severity differs between VTC-H and IP-C delivery of IOP treatment.
		PHP	Zimmerman 2021[73]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether depression symptom severity differs between VTC-H and IP-C delivery of PHP treatment.
	Telephone vs IP-C	Individual psychotherapy	Hammond 2012[70]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether depression symptom severity differs between telephone and IP-C delivery of individual psychotherapy.
Anxiety-related outcomes	VTC-H vs IP-C	Individual psychotherapy	Liou 2022[72]	Unclear	Direct	Unknown	Imprecise	Unknown	Insufficient It is unclear whether anxiety symptom severity differs between VTC-H and IP-C delivery of individual psychotherapy.
		IOP	Gannon 2021[69]	High	Direct	Unknown	Unknown	Unknown	Insufficient

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
									It is unclear whether anxiety symptom severity differs between VTC-H and IP-C delivery of IOP treatment.
	Telephone vs IP-C	Individual psychotherapy	Hammond 2012[70]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether anxiety symptom severity differs between telephone and IP-C delivery of individual psychotherapy.
Adverse events	VTC-H vs IP-C	PHP	Zimmerman 2021[73]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether the occurrence of adverse events differs between VTC-H and IP-C delivery of PHP treatment.

Abbreviations. IOP=intensive outpatient program; IP-C=clinic-based in-person; OCD=obsessive compulsive disorder; PHP=partial hospital program; PTSD=posttraumatic stress disorder; SUD=substance use disorder; VTC-C=clinic-based videoteleconference; VTC-H=home-based videoteleconference.

Strength of Evidence for KQ 2

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
PTSD									
PTSD symptom severity	VTC-H vs VTC-C	Individual psychotherapy	Franklin 2017,[11] Knowlton 2021,[56] Morland 2020[22]	Low to high	Direct	Consistent	Unknown	Unknown	Low Individual psychotherapy delivered via VTC-H may result in a similar decrease in PTSD symptom severity compared to individual psychotherapy delivered by VTC-C.

Outcome	Modality Comparison	Intervention Type	Studies	Study Limitations	Directness	Consistency	Precision	Reporting Bias	Rating and Summary of Evidence
PTSD									
PTSD diagnosis	VTC-H vs VTC-C	Individual psychotherapy	Franklin 2017[11]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether the rate of PTSD remission differs between VTC-H and VTC-C delivery of individual psychotherapy.
Depression symptom severity	VTC-H vs VTC-C	Individual psychotherapy	Franklin 2017,[11] Knowlton 2021,[56] Morland 2020[22]	Low to high	Direct	Consistent	Unknown	Unknown	Low Individual psychotherapy delivered via VTC-H may result in a similar decrease in depression symptom severity compared to individual psychotherapy delivered by VTC-C.
Anxiety symptom severity	VTC-H vs VTC-C	Individual psychotherapy	Franklin 2017[11]	High	Direct	Unknown	Unknown	Unknown	Insufficient It is unclear whether anxiety symptom severity differs between VTC-H and VTC-C delivery of individual psychotherapy.

Abbreviations. PTSD=posttraumatic stress disorder; VTC-C=clinic-based videoteleconference; VTC-H=home-based videoteleconference.

APPENDIX E: PEER REVIEW DISPOSITION

Comment #	Reviewer #	Comment	Author Response
Are the objectives, scope, and methods for this review clearly described?			
1	2	Yes	None
2	3	Yes	None
3	4	Yes	None
4	5	Yes	None
5	6	Yes	None
6	7	Yes	None
Is there any indication of bias in our synthesis of the evidence?			
7	2	No	None
8	3	No	None
9	4	No	None
10	5	No	None
11	6	No	None
12	7	No	None
Are there any published or unpublished studies that we may have overlooked?			
13	2	No	None
14	3	No	None
15	4	No	None
16	5	No	None
17	6	No	None
18	7	No	None
Additional suggestions or comments can be provided below. If applicable, please indicate the page and line numbers from the draft report.			
19	2	This report is a strong evidence brief (rapid review) about the safety and effectiveness of synchronous in-person or telehealth-based mental health care.	Thank you for this comment.

Comment #	Reviewer #	Comment	Author Response
20	2	Key Strengths include: -Thoroughness (As is typical of ESP reviews) and presentation of the studies/evidence (including describing the strength of evidence for studies reviewed and discussed) -Topic (focused and timely) -How this evidence brief builds upon other recent evidence syntheses	Thank you for this feedback.
21	2	Area for improvement: The goal (p. 8 lines 11-16) and Key question 1 (p. 9, lines 17-19) and the population eligibility criteria (p. 9 lines 30-32) all suggest that additional conditions (e.g. bipolar disorder, suicidality and/or SMI) will be reviewed and discussed. Recommend more explicitly addressing the absence of the evaluation of the studies for these conditions mentioned in KQ1.	Thank you for this feedback. We have added text to the executive summary and discussion explicitly addressing the lack of evidence available on these other conditions.
22	2	Perhaps it's as simple as editing a sentence on pg. 13, lines 9 - 13..."Most studies found....or most of the included studies...evaluated treatments for PTSD..... " and then Any of the conditions for which not enough evidence exists should be mentioned in the discussion/conclusions (e.g. as an evidence gap or area for future research).	We have added text to the literature overview stating that no studies were focused exclusively on telehealth-delivered mental health care for bipolar disorder, SMI, or suicidality. We have also added text to the future research section of the discussion more explicitly calling for research on these conditions.
23	2	Make it clearer that treatment of these other conditions were not included because of a lack of studies included after the study selection process. Since I don't have the supplementary info/search criteria, I could not tell if the terms affected what studies were found.	In the literature overview we have added text clarifying that we did not identify studies on these conditions. Search terms were included for these other conditions (see Appendix A).
24	2	Also, recommend re-wording this sentence (p. 13, line 13) ("Eight studies included participants with multiple types of mental health concerns)) to something that reflects the idea that TREATMENT was not focused on one specific condition (change the focus from who was involved to the treatment not being specific to a particular condition). For example, though the PTSD studies were the most common, clinically speaking, it is likely that those individuals also had "multiple types of mental health concerns" but those other	Thank you for this feedback. We have revised this sentence as suggested to reflect that the treatment was not focused on one specific treatment rather than the mental health concerns of the participants.

Comment #	Reviewer #	Comment	Author Response
		concerns were not the focus of treatment. The treatment was specific to one concern (PTSD). In VA, clinically speaking, most Veterans have multiple types of mental health concerns. We may just offer a treatment/ study a treatment for one concern (e.g. depression, PTSD) but there are often other issues not a focus. As you can see from the measures list Table 1, many of the PTSD studies assessed depression (e.g. with the PHQ-9, BDI=II). While I have not reviewed these studies, it is likely that many participants did have more than one MH concern (though the others may not have been a focus of treatment).	
25	3	Thank you for the opportunity to provide a review for the manuscript, “Evidence Brief: Safety and Effectiveness of Telehealth-delivered Mental Health Care” for the VA Evidence Synthesis Program. This manuscript reviewed the results of a synthesis of the literature on the safety and effectiveness of evidence based mental health treatments delivered via video teleconference technology.	Thank you for this comment.
26	3	Executive Summary and Evidence Brief: 1. Thorough and accurate summarization of the manuscript. 2. Good overview of the purpose for conducting this review, the historical context for why such a review is particularly important, and how the methodology used for this review differs from previous published reviews.	Thank you for this comment.
27	3	Method: 3. I was initially quite surprised by the evidence brief which described the low quality of research conducted in previous studies in telehealth and the lack of significance in the results in those studies. However, the methodology was described so clearly and the results so clear that the conclusions made by the end of the manuscript were well-supported.	Thank you for this comment.
28	3	Results: 4. The tables are really nicely laid out and the information made very clear for the reader.	Thank you for this comment.
29	3	5. I really appreciated the standard approach to the reporting of the results in consistent sections (symptom severity, safety,	Thank you for this comment.

Comment #	Reviewer #	Comment	Author Response
		attrition, and 'other outcomes'). This was really helpful in seeing the results normalized across studies	
30	3	Discussion: 6. The authors provided well-supported conclusions and detailed a clear direction for recommended next steps in this area of research.	Thank you for this comment.
31	3	Thank you so much for the opportunity to review this manuscript. It was enjoyable to read and provides a unique contribution to telehealth research and clinical practice.	Thank you for this comment.
32	4	My only question/suggestion is related to the title. I applaud the authors for including both RCTs as well as cohort studies, but as such, wouldn't it be more accurate to call this report "Safety, Efficacy, and Effectiveness of Telehealth-delivered Mental Health Care"? Beyond that minor suggestion, I believe this is rigorously conducted and well-written report.	Thank you for this feedback. We have carefully considered the suggested change to the report title and have decided not to make this change. Although we did include evidence from RCTs, the primary aim of the report was to examine effectiveness.
33	6	Overall, this was a very clear and well-written review. I have one major and a few minor comments.	Thank you for this comment.
34	6	Major comment 1. Throughout the document, it talks about whether each study was conducted on Veterans but does not mention whether or not it was conducted on Veterans using the VA health care system. To the extent that Veterans using VA differ systematically form Veterans not using the VA (and there should be extensive literature on this), this may be an important detail to provide.	Thank you for this feedback. We agree that this is an important point. Nearly all studies that indicated inclusion of Veteran participants were conducted within the VA health care system. We have added a section addressing this point to the discussion.
35	6	Minor comments 1. Figure 3 - The title should mention that it is depression symptom severity in PTSD	Thank you for this feedback. We have revised the Figure title to indicate that the studies in the figure are PTSD studies.
36	6	2. Document page 32, line 52 – It says "Geriatric Depression Scale (BDS)". It probably should be GDS. However, neither GDS or BDS is ever used again, so better to omit the abbreviation.	Thank you for this feedback. We did intend to use "GDS," and have removed the abbreviation.
37	6	3. Document page 47, line 39-47 – It says "Studies comparing delivery of mental health care via VTC-C versus IP-C are most relevant". This does not fit with the bottom part of the paragraph, where it says "VTC technology is now widely available and	Thank you for this feedback. We did intend to refer to VTC-H and have corrected this typo.

Comment #	Reviewer #	Comment	Author Response
		Veterans are provided access via Veteran Video Connect, potentially rendering in-clinic VTC hubs largely obsolete." Perhaps the initial sentence should be referring to VTC-H, not VTC-C?	

REFERENCES

1. Acierno R, Gros DF, Ruggiero KJ, et al. Behavioral activation and therapeutic exposure for posttraumatic stress disorder: A noninferiority trial of treatment delivered in person versus home-based telehealth. *Depression & Anxiety.* 2016;33(5):415-423.
2. Strachan M, Gros DF, Yuen E, Ruggiero KJ, Foa EB, Acierno R. Home-based telehealth to deliver evidence-based psychotherapy in veterans with ptsd. *Contemporary Clinical Trials.* 2012;33(2):402-409.
3. Acierno R, Knapp R, Tuerk P, et al. A non-inferiority trial of prolonged exposure for posttraumatic stress disorder: In person versus home-based telehealth. *Behaviour Research & Therapy.* 2017;89:57-65.
4. Gros DF, Allan NP, Lancaster CL, Szafranski DD, Acierno R. Predictors of treatment discontinuation during prolonged exposure for ptsd. *Behavioural & Cognitive Psychotherapy.* 2018;46(1):35-49.
5. Gros DF, Lancaster CL, Lopez CM, Acierno R. Treatment satisfaction of home-based telehealth versus in-person delivery of prolonged exposure for combat-related ptsd in veterans. *Journal of Telemedicine & Telecare.* 2018;24(1):51-55.
6. Reich K, Nemeth LS, Mueller M, Sternke LM, Acierno R. Does psychosocial functioning improve with prolonged exposure in veterans with ptsd?: Exploring traditional and home-based telehealth delivery methods. *Journal of Psychosocial Nursing & Mental Health Services.* 2021;59(2):31-40.
7. Yuen EK, Gros DF, Price M, et al. Randomized controlled trial of home-based telehealth versus in-person prolonged exposure for combat-related ptsd in veterans: Preliminary results. *Journal of Clinical Psychology.* 2015;71(6):500-512.
8. Acierno R, Jaffe AE, Gilmore AK, et al. A randomized clinical trial of in-person vs. Home-based telemedicine delivery of prolonged exposure for ptsd in military sexual trauma survivors. *Journal of Anxiety Disorders.* 2021;83:102461.
9. Gilmore AK, Lopez C, Muzzy W, et al. Emotion dysregulation predicts dropout from prolonged exposure treatment among women veterans with military sexual trauma-related posttraumatic stress disorder. *Womens Health Issues.* 2020;30(6):462-469.
10. Lopez CM, Gilmore AK, Brown WJ, et al. Effects of emotion dysregulation on post-treatment post-traumatic stress disorder and depressive symptoms among women veterans with military sexual trauma. *Journal of Interpersonal Violence.* 2021:8862605211005134.
11. Franklin CL, Cuccurullo LA, Walton JL, Arseneau JR, Petersen NJ. Face to face but not in the same place: A pilot study of prolonged exposure therapy. *Journal of Trauma & Dissociation.* 2017;18(1):116-130.
12. Frueh BC, Monnier J, Grubaugh AL, Elhai JD, Yim E, Knapp R. Therapist adherence and competence with manualized cognitive-behavioral therapy for ptsd delivered via videoconferencing technology. *Behavior Modification.* 2007;31(6):856-866.
13. Frueh BC, Monnier J, Yim E, Grubaugh AL, Hamner MB, Knapp RG. A randomized trial of telepsychiatry for post-traumatic stress disorder. *Journal of Telemedicine & Telecare.* 2007;13(3):142-147.
14. Glassman LH, Mackintosh MA, Talkovsky A, et al. Quality of life following treatment for ptsd: Comparison of videoconferencing and in-person modalities. *Journal of Telemedicine & Telecare.* 2019;25(2):123-127.

15. Haghnia Y, Samad-Soltani T, Yousefi M, Sadr H, Rezaei-Hachesu P. Telepsychiatry-based care for the treatment follow-up of iranian war veterans with post-traumatic stress disorder: A randomized controlled trial. *Iranian Journal of Medical Sciences.* 2019;44(4):291-298.
16. Hernandez-Tejada MA, Zoller JS, Ruggiero KJ, Kazley AS, Acierno R. Early treatment withdrawal from evidence-based psychotherapy for ptsd: Telemedicine and in-person parameters. *International Journal of Psychiatry in Medicine.* 2014;48(1):33-55.
17. Liu L, Thorp SR, Moreno L, et al. Videoconferencing psychotherapy for veterans with ptsd: Results from a randomized controlled non-inferiority trial. *Journal of Telemedicine & Telecare.* 2020;26(9):507-519.
18. Maieritsch KP, Smith TL, Hessinger JD, Ahearn EP, Eickhoff JC, Zhao Q. Randomized controlled equivalence trial comparing videoconference and in person delivery of cognitive processing therapy for ptsd. *Journal of Telemedicine & Telecare.* 2016;22(4):238-243.
19. Morland LA, Mackintosh M-A, Greene CJ, et al. Cognitive processing therapy for posttraumatic stress disorder delivered to rural veterans via telemental health: A randomized noninferiority clinical trial. *The Journal of Clinical Psychiatry.* 2014;75(5):470-476.
20. Morland LA, Hynes AK, Mackintosh M-A, Resick PA, Chard KM. Group cognitive processing therapy delivered to veterans via telehealth: A pilot cohort. *Journal of Traumatic Stress.* 2011;24(4):465-469.
21. Morland LA, Mackintosh MA, Rosen CS, et al. Telemedicine versus in-person delivery of cognitive processing therapy for women with posttraumatic stress disorder: A randomized noninferiority trial. *Depression & Anxiety.* 2015;32(11):811-820.
22. Morland LA, Mackintosh M-A, Glassman LH, et al. Home-based delivery of variable length prolonged exposure therapy: A comparison of clinical efficacy between service modalities. *Depression and Anxiety.* 2020;37(4):346-355.
23. Morland LA, Knopp KC, Khalifian CE, et al. A randomized trial of brief couple therapy for ptsd and relationship satisfaction. *Journal of Consulting and Clinical Psychology.* 2022;90(5):392-404.
24. Peterson AL, Mintz J, Moring JC, et al. In-office, in-home, and telehealth cognitive processing therapy for posttraumatic stress disorder in veterans: A randomized clinical trial. *BMC Psychiatry.* 2022;22(1):41.
25. White CN, Kauffman BY, Acierno R. Factors contributing to veterans' satisfaction with ptsd treatment delivered in person compared to telehealth. *Journal of Telemedicine & Telecare.* 2021:1357633X20987704.
26. Ziemba SJ, Bradley NS, Landry L-AP, Roth CH, Porter LS, Cuyler RN. Posttraumatic stress disorder treatment for operation enduring freedom/operation iraqi freedom combat veterans through a civilian community-based telemedicine network. *Telemedicine and e Health.* 2014;20(5):446-450.
27. Alegria M, Ludman E, Kafali EN, et al. Effectiveness of the engagement and counseling for latinos (ecla) intervention in low-income latinos. *Medical Care.* 2014;52(11):989-997.
28. Aguilera A, Ramos Z, Sistiva D, Wang Y, Alegria M. Homework completion via telephone and in-person cognitive behavioral therapy among latinos. *Cognitive Therapy & Research.* 2018;42(3):340-347.
29. Alcantara C, Li X, Wang Y, Canino G, Alegria M. Treatment moderators and effectiveness of engagement and counseling for latinos intervention on worry reduction

in a low-income primary care sample. *Journal of Consulting & Clinical Psychology.* 2016;84(11):1016-1022.
30. Kafali N, Cook B, Canino G, Alegria M. Cost-effectiveness of a randomized trial to treat depression among latinos. *The Journal of Mental Health Policy & Economics.* 2014;17(2):41-50.
31. Choi NG, Hegel MT, Marti N, Marinucci ML, Sirrianni L, Bruce ML. Telehealth problem-solving therapy for depressed low-income homebound older adults. *American Journal of Geriatric Psychiatry.* 2014;22(3):263-271.
32. Egede LE, Acierno R, Knapp RG, et al. Psychotherapy for depression in older veterans via telemedicine: A randomised, open-label, non-inferiority trial. *The Lancet Psychiatry.* 2015;2(8):693-701.
33. 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. *Journal of Clinical Psychiatry.* 2016;77(12):1704-1711.
34. Egede LE, Gebregziabher M, Walker RJ, Payne EH, Acierno R, Frueh BC. Trajectory of cost overtime after psychotherapy for depression in older veterans via telemedicine. *Journal of Affective Disorders.* 2017;207:157-162.
35. Egede LE, Dismuke CE, Walker RJ, Acierno R, Frueh BC. Cost-effectiveness of behavioral activation for depression in older adult veterans: In-person care versus telehealth. *Journal of Clinical Psychiatry.* 2018;79(5):28.
36. Hungerbuehler I, Valiengo L, Loch AA, Rossler W, Gattaz WF. Home-based psychiatric outpatient care through videoconferencing for depression: A randomized controlled follow-up trial. *JMIR Mental Health.* 2016;3(3):e36.
37. 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. *Journal of Consulting & Clinical Psychology.* 2016;84(11):923-934.
38. Bounthavong M, Pruitt LD, Smolenski DJ, Gahm GA, Bansal A, Hansen RN. Economic evaluation of home-based telebehavioural health care compared to in-person treatment delivery for depression. *Journal of telemedicine and telecare.* 2018;24(2):84-92.
39. Pruitt LD, Vuletic S, Smolenski DJ, Wagner A, Luxton DD, Gahm GA. Predicting post treatment client satisfaction between behavioural activation for depression delivered either in-person or via home-based telehealth. *Journal of Telemedicine & Telecare.* 2019;25(8):460-467.
40. Smolenski DJ, Pruitt LD, Vuletic S, Luxton DD, Gahm G. Unobserved heterogeneity in response to treatment for depression through videoconference. *Psychiatr Rehabil J.* 2017;40(3):303-308.
41. Mohr DC, Ho J, Duffecy J, et al. Effect of telephone-administered vs face-to-face cognitive behavioral therapy on adherence to therapy and depression outcomes among primary care patients: A randomized trial. *JAMA.* 2012;307(21):2278-2285.
42. Kalapatapu RK, Ho J, Cai X, Vinogradov S, Batki SL, Mohr DC. Cognitive-behavioral therapy in depressed primary care patients with co-occurring problematic alcohol use: Effect of telephone-administered vs. Face-to-face treatment-a secondary analysis. *Journal of Psychoactive Drugs.* 2014;46(2):85-92.
43. Stiles-Shields C, Kwasny MJ, Cai X, Mohr DC. Therapeutic alliance in face-to-face and telephone-administered cognitive behavioral therapy. *Journal of Consulting & Clinical Psychology.* 2014;82(2):349-354.

44. Stiles-Shields C, Corden ME, Kwasny MJ, Schueller SM, Mohr DC. Predictors of outcome for telephone and face-to-face administered cognitive behavioral therapy for depression. *Psychological Medicine.* 2015;45(15):3205-3215.
45. Ruskin PE, Silver-Aylaian M, Kling MA, et al. Treatment outcomes in depression: Comparison of remote treatment through telepsychiatry to in-person treatment. *American Journal of Psychiatry.* 2004;161(8):1471-1476.
46. Rhein AN. The effect of personality disorder characteristics on live versus remote treatment for veterans with depressive disorder. *Dissertation Abstracts International: Section B: The Sciences and Engineering.* 2001;62(2-B).
47. Cherestal S. Remotely conducted versus office-based virtual reality treatment for fear of flying. *Dissertation Abstracts International: Section B: The Sciences and Engineering.* 2019;80(11-B(E).
48. Lovell K, Cox D, Haddock G, et al. Telephone administered cognitive behaviour therapy for treatment of obsessive compulsive disorder: Randomised controlled non-inferiority trial. *BMJ.* 2006;333(7574):883.
49. Watts S, Marchand A, Bouchard S, et al. Telepsychotherapy for generalized anxiety disorder: Impact on the working alliance. *Journal of Psychotherapy Integration.* 2020;30(2):208-225.
50. Cuevas CDL, Arredondo MT, Cabrera MF, Sulzenbacher H, Meise U. Randomized clinical trial of telepsychiatry through videoconference versus face-to-face conventional psychiatric treatment. *Telemedicine Journal & e-Health.* 2006;12(3):341-350.
51. Richard O'Reilly MB, F.R.C.P.C. ,, Joan Bishop MD, F.R.C.P.C. ,, Karen Maddox RN, M.A. ,, Lois Hutchinson MD, F.R.C.P.C. ,, Michael Fisman MB, F.R.C.P.C. ,, Jatinder Takhar MD, F.R.C.P.C. ,. Is telepsychiatry equivalent to face-to-face psychiatry? Results from a randomized controlled equivalence trial. *Psychiatric Services.* 2007;58(6):836-843.
52. Stubbings DR, Rees CS, Roberts LD, Kane RT. Comparing in-person to videoconference-based cognitive behavioral therapy for mood and anxiety disorders: Randomized controlled trial. *Journal of Medical Internet Research.* 2013;15(11):e258.
53. Germain V, Marchand A, Bouchard S, Drouin MS, Guay S. Effectiveness of cognitive behavioural therapy administered by videoconference for posttraumatic stress disorder. *Cognitive Behaviour Therapy.* 2009;38(1):42-53.
54. Germain V, Marchand A, Bouchard S, Guay S, Drouin MS. Assessment of the therapeutic alliance in face-to-face or videoconference treatment for posttraumatic stress disorder. *Cyberpsychology, behavior and social networking.* 2010;13(1):29-35.
55. Gros DF, Yoder M, Tuerk PW, Lozano BE, Acierno R. Exposure therapy for ptsd delivered to veterans via telehealth: Predictors of treatment completion and outcome and comparison to treatment delivered in person. *Behavior Therapy.* 2011;42(2):276-283.
56. Knowlton CN, Nelson KG. Ptsd telehealth treatments for veterans: Comparing outcomes from in-person, clinic-to-clinic, and home-based telehealth therapies. *Journal of Rural Mental Health.* 2021;45(4):243-255.
57. LoSavio ST, Worley CB, Aajmain ST, Rosen CS, Wiltsey Stirman S, Sloan DM. Effectiveness of written exposure therapy for posttraumatic stress disorder in the department of veterans affairs health care system. *Psychological Trauma:Theory, Pesearch, Practice and Policy.* 2021;28:28.
58. Tuerk PW, Yoder M, Ruggiero KJ, Gros DE, Acierno R. A pilot study of prolonged exposure therapy for posttraumatic stress disorder delivered via telehealth technology. *Journal of Traumatic Stress.* 2010;23(1):116-123.

59. Valentine LM, Donofry SD, Broman RB, Smith ER, Rauch SA, Sexton MB. Comparing ptsd treatment retention among survivors of military sexual trauma utilizing clinical video technology and in-person approaches. *Journal of Telemedicine & Telecare.* 2020;26(7-8):443-451.
60. Wierwille JL, Pukay-Martin ND, Chard KM, Klump MC. Effectiveness of ptsd telehealth treatment in a va clinical sample. *Psychological Services.* 2016;13(4):373-379.
61. Ritchie LG. Group telepsychotherapy for reducing depression: An exploratory study. *Dissertation Abstracts International: Section B: The Sciences and Engineering.* 2008;68(11-B).
62. Bouchard S, Paquin B, Payeur R, et al. Delivering cognitive-behavior therapy for panic disorder with agoraphobia in videoconference. *Telemedicine Journal & E-Health.* 2004;10(1):13-25.
63. Bouchard S, Allard M, Robillard G, et al. Videoconferencing psychotherapy for panic disorder and agoraphobia: Outcome and treatment processes from a non-randomized non-inferiority trial. *Frontiers in Psychology.* 2020;11:2164.
64. Milosevic I, Cameron DH, Milanovic M, McCabe RE, Rowa K. Face-to-face versus video teleconference group cognitive behavioural therapy for anxiety and related disorders: A preliminary comparison. *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie.* 2022;67(5):391-402.
65. Pinciotti CM, Bulkes NZ, Horvath G, Riemann BC. Efficacy of intensive cbt telehealth for obsessive-compulsive disorder during the covid-19 pandemic. *Journal of Obsessive-Compulsive & Related Disorders.* 2022;32:100705.
66. Bean CAL, Aurora P, Maddox CJ, Mekota R, Updegraff A. A comparison of telehealth versus in-person group therapy: Results from a dbt-based dual diagnosis iop. *Journal of Clinical Psychology.* 2022;09:09.
67. Vakkalanka JP, Lund BC, Ward MM, et al. Telehealth utilization is associated with lower risk of discontinuation of buprenorphine: A retrospective cohort study of us veterans. *Journal of General Internal Medicine.* 2022;37(7):1610-1618.
68. Zheng W, Nickasch M, Lander L, et al. Treatment outcome comparison between telepsychiatry and face-to-face buprenorphine medication-assisted treatment for opioid use disorder: A 2-year retrospective data analysis. *Journal of Addiction Medicine.* 2017;11(2):138-144.
69. Gannon JM, Brar JS, Zawacki S, Painter T, O'Toole K, Chengappa KNR. From office-based treatment to telehealth: Comparing clinical outcomes and patient participation in a psychiatric intensive outpatient program with a large transdiagnostic sample. *Telemedicine Journal & E Health.* 2021;29:29.
70. Hammond GC, Croudace TJ, Radhakrishnan M, et al. Comparative effectiveness of cognitive therapies delivered face-to-face or over the telephone: An observational study using propensity methods. *PLoS ONE [Electronic Resource].* 2012;7(9):e42916.
71. Khatri N, Marziali E, Tchernikov I, Shepherd N. Comparing telehealth-based and clinic-based group cognitive behavioral therapy for adults with depression and anxiety: A pilot study. *Clinical Interventions In Aging.* 2014;9:765-770.
72. Liou H, Lane C, Huang C, Mookadam M, Joseph M, Hecker DuVal J. Eye movement desensitization and reprocessing in a primary care setting: Assessing utility and comparing efficacy of virtual versus in-person methods. *Telemedicine Journal & E Health.* 2022;08:08.
73. Zimmerman M, Terrill D, D'Avanzato C, Tirpak JW. Telehealth treatment of patients in an intensive acute care psychiatric setting during the covid-19 pandemic: Comparative

safety and effectiveness to in-person treatment. *Journal of Clinical Psychiatry.* 2021;82(2):16.
74. Cherestal S, Schare ML, Grimaldi S, Costello K, DeJesus C. Remotely conducted versus office-based virtual reality treatment for aviophobia: Questions of feasibility and accessibility. *Translational Issues in Psychological Science.* 2021;7(3):218-228.
75. Cherestal S. *Remotely conducted versus office-based virtual reality treatment for fear of flying.* Hofstra University; 2019.
76. Choi NG, Hegel MT, Marti CN, Marinucci ML, Sirrianni L, Bruce ML. Telehealth problem-solving therapy for depressed low-income homebound older adults. *The American Journal of Geriatric Psychiatry.* 2014;22(3):263-271.
77. Germain V, Marchand A, Bouchard S, Drouin M-S, Guay S. Effectiveness of cognitive behavioural therapy administered by videoconference for posttraumatic stress disorder. *Cognitive Behaviour Therapy.* 2009;38(1):42-53.