Comparison of Mental Health Symptom Changes from pre-COVID-19 to COVID-19 by Sex or Gender: A Systematic Review and Meta-analysis

Tiffany Dal Santo[1,2]; Ying Sun, MPH[1]; Yin Wu, PhD[1,3]; Chen He, MScPH[1]; Yutong Wang[1]; Xiaowen Jiang, BA[1]; Kexin Li, BSc[1]; Olivia Bonardi[1]; Ankur Krishnan, MSc[1]; Jill T. Boruff, MLIS[4]; Danielle B. Rice, MSc[1,2]; Sarah Markham, PhD[5]; Brooke Levis, PhD[6]; Marleine Azar, MSc[1]; Dipika Neupane, MSc[1]; Amina Tasleem[1,2]; Anneke Yao[1,2]; Ian Thombs-Vite[1]; Branka Agic, MD[7,8]; Christine Fahim, PhD[9]; Michael S. Martin, PhD[10,11]; Sanjeev Sockalingam, MD[7,12]; Gustavo Turecki, MD[2,13]; Andrea Benedetti, PhD[14-16]; Brett D. Thombs, PhD[1-3,14,15,17,18]

[1]Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada; [2]Department of Psychology, McGill University, Montréal, Québec, Canada; [3]Department of Psychiatry, McGill University, Montréal, Québec, Canada; [4]Schulich Library of Physical Sciences, Life Sciences, and Engineering, McGill University, Montréal, Québec, Canada; [5]Department of Biostatistics and Health Informatics, King's College London, London, UK; [6]Centre for Prognosis Research, School of Medicine, Keele University, Staffordshire, UK; [7]Centre for Addiction and Mental Health, Toronto, Ontario, Canada; [8]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; [9]Li Ka Shing Knowledge Institute, Unity Health Toronto, Toronto, Ontario, Canada; [10]School of Epidemiology and Public Health, University of Ottawa; Ontario, Canada; [11]Correctional Service of Canada, Ottawa, Ontario, Canada; [12]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada; [13]McGill Group for Suicide Studies, Douglas Mental Health University Institute, McGill University, Montreal, Quebec, Canada; [14]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, Québec, Canada; [15]Department of Medicine, McGill University, Montréal, Québec, Canada; [16]Respiratory Epidemiology and Clinical Research Unit, McGill University Health Centre, Montréal, Québec, Canada; [17]Department of

Educational and Counselling Psychology, McGill University, Montréal, Québec, Canada;

[18]Biomedical Ethics Unit, McGill University, Montréal, Québec, Canada.

Corresponding author:

Brett D. Thombs, PhD; Jewish General Hospital; 4333 Cote Ste Catherine Road; Montreal, Quebec, Canada, H3T 1E4; Tel (514) 340-8222 ext. 25112; E-mail: brett.thombs@mcgill.ca

ORCID: 0000-0002-5644-8432

Word count: 2,499

Contributions: YS, YWu, DBR, AB, and BDT were responsible for the study conception and design. JTB was responsible for the design of the database searches. AK carried out the searches. TDS, YS, YWu, CH, YWang, XJ, KL, OB, AK, DBR, SM, MA, DN, AT, AY, ITV and BDT contributed to data extraction, coding, and evaluation of included studies. YS, OB, and CH were responsible for study coordination. TDS, YS, YWu, BL, AB, and BDT were involved in data analysis. BA, CF, MSM, SS, and GT contributed to interpretation of results as knowledge translation partners. TDS and BDT drafted the manuscript. All authors provided a critical review and approved the final manuscript. BDT is the guarantor; he had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analyses. BDT is the corresponding author and attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: The study was funded by the Canadian Institutes of Health Research (CMS-171703; MS1-173070) and McGill Interdisciplinary Initiative in Infection and Immunity Emergency COVID-19 Research Fund (R2-42). YWu and BL were supported by Fonds de recherche du Québec – Santé (FRQS) Postdoctoral Training Fellowships. DBR was supported by a Vanier Canada Graduate Scholarship. AB was supported by a FRQS senior researcher salary award. BDT was supported by a Tier 1 Canada Research Chair.

Declaration of Competing Interests: All authors have completed the ICJME uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years. All authors declare no relationships or activities that could appear to have influenced the submitted work. No funder had any role in the design and conduct of the study; collection, management, analysis, and

interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Sharing: All data used in the study are available in the manuscript and its tables or online at https://www.depressd.ca/covid-19-mental-health.

ABSTRACT

Background: Women and gender-diverse individuals face disproportionate socioeconomic burden during COVID-19. We compared mental health symptom changes since pre-COVID-19 by sex or gender.

Methods: We searched MEDLINE, PsycINFO, CINAHL, EMBASE, Web of Science, China National Knowledge Infrastructure, Wanfang, medRxiv, and Open Science Framework December 31, 2019 to March 22, 2021 for studies that reported mental health outcomes prior to and during COVID-19 by sex or gender. We conducted restricted maximum-likelihood random-effects meta-analyses.

Results: All 11 included studies (9 unique cohorts) compared females or women to males or men; none included gender-diverse individuals. Continuous symptom change differences were not statistically significant for depression (standardized mean difference [SMD]= 0.15, 95% CI -0.09 to 0.39; 3 studies, 4,159 participants; I^2=77%) and stress (SMD= -0.09, 95% CI -0.21 to 0.02; 3 studies, 1,217 participants; I^2=0%), but anxiety (SMD= 0.14, 95% CI 0.01 to 0.26; 3 studies, 4,028 participants; I^2=34%) and general mental health (SMD= 0.15, 95% CI 0.12 to 0.18; 2 studies, 15,590 participants; I^2=0%) worsened more among females or women than males or men. There were no significant differences in changes in proportion above a cut-off: anxiety (difference= 0.00, 95% CI -0.01 to 0.02; 2 studies, 6,684 participants; I^2=0%), depression (difference= 0.12, 95% CI -0.04 to 0.28; 1 study, 217 participants), general mental health (difference= -0.03, 95% CI -0.09 to 0.04; 3 studies, 18,985 participants; I^2=94%), stress (difference= 0.04, 95% CI -0.11 to 0.18; 1 study, 217 participants).

Interpretation: Mental health outcomes did not differ or were somewhat worse among women than men.

The COVID-19 pandemic has caused over 3.8 million deaths and disrupted social and economic activities across the globe.[1,2] Men infected with COVID-19 are at greater risk of intensive care admission and death than women,[3,4] but socioeconomic burden has disproportionately impacted women.[5-12] Economically, most single parents are women, and women earn less, are more likely to live in poverty, and hold less secure jobs than men, which heightens vulnerability during COVID-19.[5,8-11] Women are overrepresented in health care, which involves infection risk.[5-10] They provide most childcare and family elder care.[5,8-10] Intimate partner violence has increased with the majority directed towards women.[5,7-10,12] Sex and gender minority individuals may also face additional socioeconomic challenges during COVID-19.[13,14]

Some studies have reported that COVID-19 mental health effects have been greater for women than men.[5,15-18] These, however, have been cross-sectional studies that evaluated proportions of participants above cut-offs on self-report measures without consideration of pre-COVID-19 differences, even though mental health disorders and symptoms are more common among women.[19-23]

Evidence from longitudinal cohorts that compare mental health symptoms pre-COVID-19 to during COVID-19 is needed. We are conducting a series of living systematic reviews on COVID-19 mental health,[24-26] including mental health changes.[26] The objective of this study was to compare mental health changes by sex or gender.

METHODS

Our series of living systematic review was registered in PROSPERO (CRD42020179703), and a protocol was posted pre-initiation (https://osf.io/96csg/). The present study is a sub-study of our main mental health changes review.[26] Results are reported per the PRISMA statement.[27]

Study Eligibility

For our main symptom changes review, studies on any population were included if they compared mental health outcomes assessed between January 1, 2018 and December 31,

2019, when China first reported COVID-19 to the World Health Organization,[28] to outcomes collected January 1, 2020 or later. Compared samples had to include at least 90% of the same participants pre-COVID-19 and during COVID-19 or use statistical methods to account for missing data. Studies with < 100 participants were excluded for feasibility and due to their limited relative value. For the present analysis, studies had to report mental health outcomes separately by sex (assignment based on external genitalia, usually at birth; e.g., female, male, intersex) or gender (socially constructed characteristics of roles and behaviours; e.g., woman, man, trans woman, trans man, non-binary).[29]

Search Strategy

MEDLINE (Ovid), PsycINFO (Ovid), CINAHL (EBSCO), EMBASE (Ovid), Web of Science Core Collection: Citation Indexes, China National Knowledge Infrastructure, Wanfang, medRxiv (preprints), and Open Science Framework Preprints (preprint server aggregator) were searched using a strategy designed by an experienced health science librarian. The China National Knowledge Infrastructure and Wanfang databases were searched using Chinese terms based on our English-language strategy. The rapid project launch did not allow for formal peer review, but COVID-19 terms were developed in collaboration with other librarians working on the topic. See Appendix 1 for search strategies. The initial search was conducted from December 31, 2019 to April 13, 2020 with automated daily updates. We converted to weekly updates on December 28, 2020 to increase processing efficiency.

Selection of Eligible Studies

Search results were uploaded into DistillerSR (Evidence Partners, Ottawa, Canada). Duplicate references were removed. Then two reviewers independently evaluated titles and abstracts in random order; if either reviewer believed a study was potentially eligible, it underwent full-text review by two independent reviewers. Discrepancies at the full-text level were resolved by consensus, with a third reviewer consulted if necessary. An inclusion and

exclusion coding guide was developed, and team members were trained over several sessions. See Appendix 2.

Data Extraction

For each eligible study, data were extracted by a single reviewer using a pre-specified form with validation by a second reviewer. Reviewers extracted (1) publication characteristics (e.g., first author, year, journal); (2) population characteristics and demographics, including eligibility criteria, recruitment method, number of participants, assessment timing, age; (3) mental health outcomes, which were defined broadly and could have included, for example, symptoms of anxiety, depression, stress, loneliness, anger, grief, or other emotional disturbance; (4) if studies reported outcomes by sex or gender or used these terms inconsistently (e.g., described using gender but reported results for females and males, which are sex terms); and (5) if sex or gender were treated as binary or categorical.

Adequacy of study methods and reporting was assessed using an adapted version of the Joanna Briggs Institute Checklist for Prevalence Studies, which assesses appropriateness of the sampling frame for the target population, appropriateness of recruiting methods, sample size, description of setting and participants, participation or response rate, outcome assessment methods, standardization of assessments across participants, appropriateness of statistical analyses, and follow-up.[30] See Appendix 3. Discrepancies were resolved between reviewers with a third reviewer consulted if necessary.

Statistical Analyses

For changes from pre-COVID-19 to COVID-19, in continuous and dichotomous outcomes, within sex or gender groups, we presented results as positive when mental health worsened pre-COVID-19 to COVID-19 and negative when it improved. For continuous outcomes, separately for each sex or gender group, we extracted a standardized mean difference (SMD) effect size with 95% confidence intervals (CIs) for change from pre-COVID-19 to COVID-19. If not provided, we calculated it using Hedges' g,[31] as described by Borenstein et

al.[32] For each study, we then calculated a Hedges' g difference in change between sex or gender groups with 95% CI.

For studies that reported proportions of participants above a scale cut-off, for pre-COVID-19 and COVID-19 proportions, if not provided, we calculated a 95% CI using Agresti and Coull's approximate method for binomial proportions.[33] We then extracted or calculated the proportion change in participants above the cut-off, along with 95% CI, for each sex or gender group. If 95% CIs were not reported, we generated them using Newton's method for differences between binomial proportions based on paired data.[34] To do this, which requires the number of cases at both assessments, which is not typically available, we assumed that 50% of pre-COVID-19 cases continued to be cases during COVID-19 and confirmed that results did not differ substantively if we used values from 30% to 70% (all 95% CI end points within 0.02; see Appendix 4). Finally, we calculated a difference of the proportion change between sex or gender groups with 95% CI.[35]

Meta-analyses were done to synthesize differences between sex or gender groups in SMD change for continuous outcomes and in proportion change for dichotomous outcomes via restricted maximum-likelihood random-effects meta-analysis. Heterogeneity was assessed with the I^2 statistic. Meta-analysis was performed in R (R version 3.6.3, RStudio Version 1.2.5042), using the metacont and metagen functions in the meta package.[36] Forest plots were generated using the forest function in meta. Positive values indicated more relatively worse changes in mental health for females or women compared to males or men.

RESULTS

Search Results and Selection of Eligible Studies

As of March 22, 2021, there were 45,777 unique references identified and screened for potential eligibility, of which 45,251 were excluded after title and abstract review and 394 after full-text review. Of 132 remaining articles, 121 were excluded because they compared symptoms at different points during the pandemic but not to pre-COVID-19 data or because

they did not compare results by sex or gender, leaving 11 included studies that reported data from 9 cohorts (Figure 1).

Characteristics of Included Studies

Four publications[37-40] reported on two large, national, probability-based samples from the United Kingdom (N = 10,918 to 15,376)[37,38] and the Netherlands (N = 3,983 to 4,064).[39,40] One study[41] reported on adults from a national sample from China (N = 6,467). Two studies[42,43] assessed young adults; one reported on a sample of twins from the United Kingdom (N = 3,563)[42] and another on a sample from Switzerland (N = 786).[43] One study assessed adolescents from Australia (N = 248),[44] and three studies[45-47] assessed undergraduate students from China (N = 4,085),[45] India (N = 217),[46] and the United Kingdom (N = 214).[47] Four studies assessed anxiety symptoms,[41,42,44,46] 3 depression symptoms,[42,44,46] 4 (2 cohorts) general mental health,[37-40] and 3 stress.[43,46,47] Table 1 shows study characteristics. All studies compared women and men or females and males; none included other sex or gender groups. Results during COVID-19 were assessed between March and June 2020 for all studies. Two cohorts also reported results from September 2020[38] and November to December 2020.[40]

Adequacy of Study Methods and Reporting

Four studies (2 cohorts)[37-40] were rated as “yes” for adequacy for all items[39,40] or “yes” on all items except one, which was rated “unclear”.[37,38] Other studies were rated “no” for 1-3 items (plus 0-2 unclear ratings)[41,43-47] or “no” on none but unclear on 3 items.[42] There were 5 studies[43-47] rated “no” for appropriate sampling frame (45%), 7 “no” or “unclear” for adequate response rate and coverage (64%),[37,38,41-44,47] and 4 “no” or “unclear” for follow-up response rate and management (36%).[41,42,44,47] See Table 2.

Mental Health Symptom Changes

There was a total of 11 comparisons of continuous score changes and 9 of proportion changes; in all but one, females or women had higher scores or proportions above a cut-off pre-COVID-19. Mental health scores and symptom changes for all outcome domains are reported

separately by sex or gender groups in Table 3. Differences in continuous and dichotomous changes by sex or gender are shown in Figures 2 to 5. Estimates of difference in change by sex or gender were close to zero and not statistically significant for anxiety symptoms with dichotomous outcomes (Figure 2b; proportion change difference = 0.00, 95% CI -0.01 to 0.02; N = 2 studies,[41,46] 6,684 participants; I^2 = 0%), depression symptoms with continuous (Figure 3a; SMD change difference = 0.15, 95% CI -0.09 to 0.39; N = 3 studies,[42,44,46] 4,159 participants; I^2 = 77%) and dichotomous outcomes (Figure 3b; proportion change difference = 0.12, 95% CI -0.04 to 0.28; N = 1 study,[46] 217 participants), general mental health dichotomous outcomes (Figure 4b [all results from early 2020]; proportion change difference = -0.03, 95% CI -0.09 to 0.04; N = 3 studies,[38,39,45] 18,985 participants; I^2 = 94%), and stress with continuous (Figure 5a; SMD change difference = -0.09, 95% CI -0.21 to 0.02; N = 3 studies,[43,46,47] 1,217 participants; I^2 = 0%) and dichotomous outcomes (Figure 5b; proportion change difference = 0.04, 95% CI -0.11 to 0.18; N = 1 study,[46] 217 participants). Of the 3 studies that reported dichotomous general mental health, 2 also reported outcomes from late 2020; when those results were used, the null finding did not change (Figure 4c; proportion change difference = 0.00, 95% CI -0.02 to 0.03; N = 3 studies,[38,40,45] 19,067 participants; I^2 = 67%).

Anxiety, measured continuously, worsened significantly more for females or women than for males or men (Figure 2a; SMD change difference = 0.14, 95% CI 0.01 to 0.26; N = 3 studies,[42,44,46] 4,028 participants; I^2= 34%). General mental health, measured continuously, also worsened more for females or women than for males or men in early COVID-19 (Figure 4a; SMD difference in change = 0.15, 95% CI 0.12 to 0.18; N = 2 studies,[37,47] 15,590 participants; I^2 = 0%). This was predominantly based on a large population-based study from the United Kingdom.[37] That study did not report results from fall 2020 for continuous outcomes, but as shown in Table 3 and Figures 4b and 4c, the difference in change between females or women and males or men decreased between early and late 2020 for dichotomous outcomes in the

same cohort.[38] The magnitude of both statistically significant differences was small (see Figure 6).

DISCUSSION

The COVID-19 pandemic has affected women and gender minorities disproportionately.[5-14] We reviewed evidence from studies that reported mental health changes from pre-COVID-19 to COVID-19 by sex or gender. We examined 11 studies (9 cohorts) that reported on anxiety symptoms, depression symptoms, general mental health, and stress. We compared females or women with males or men; no studies compared gender minorities with any other group. Syntheses of continuously measured anxiety symptoms (SMD = 0.14, 95% CI 0.01 to 0.26) and general mental health (SMD = 0.15, 95% CI 0.12 to 0.18) found that mental health worsened more for females or women than males or men, but the magnitude was small and not typically considered clinically important.[48] No other mental health outcomes changed more for females or women than males or men.

Sex and gender differences in mental health disorder prevalence, symptoms, and risk factors are well-established.[49-52] Likely risk factors include gender inequities and discrimination, higher rates of interpersonal stressors, and violence,[53,54] and many of these risk factors have been exacerbated for women during COVID-19.[5-12] Overall mental health does not appear to have changed substantively for the general population based on our main systematic review,[26] but there is concern that there could, nonetheless, be important sex- or gender-based differences. We found small symptom change differences that reflect relative worsening for women for 2 of 8 outcomes assessed, but no differences appeared to be clinically meaningful.

Based on our findings, it is possible that despite the challenges women have faced, most have been resilient and that the mental health disaster that has been predicted by many has not occurred.[55] This finding departs from what has been reported in some research and by the media. Three factors may feed this discrepancy. One is the publication of many cross-sectional studies that report proportions above cut-offs on self-report measures, which are not designed

for that purpose, and assume that high numbers must not have been present pre-COVID-19.[26] A second is the use of surveys that ask questions about well-being with COVID-19 explicitly assigned as a cause; illustrating the pitfalls of this, a study of over 2,000 young Swiss adult men found significant angst when questions were asked in this way, but no changes in validated measures of depression symptoms and stress from pre-COVID.[56] A third reason relates to news media reports that emphasize dramatic events and anecdotes without evidence that demonstrates changes.[55]

Strengths of our study include the use of rigorous systematic review methods; searching 9 databases, including Chinese-language databases, without language restrictions; and including studies that enabled the direct comparison of mental health changes by sex or gender. There are limitations to consider. First, this review only included 11 studies from 9 cohorts, and many had limitations related to study sampling frames and recruitment methods, follow-up rates, and management of missing data. Second, heterogeneity was high for some meta-analyses. Third, there were not enough studies to attempt sub-group analyses by additional sociodemographic or other factors. Fourth, we did not identify any studies that compared results from gender-diverse individuals to other gender groups.

In sum, we identified small sex- or gender-based differences for anxiety symptoms and general mental health, continuously measured, but other outcomes (continuous depression symptoms and stress; dichotomous anxiety symptoms, depression symptoms, general mental health, stress) were not different by sex or gender. These are aggregate results, though, and many individuals have certainly experienced negative mental health changes related to increased socioeconomic burden. Mental health changes should continue to be monitored in COVID-19, taking into consideration sex and gender, and studies should examine reasons for what appears to be resilience among many women despite facing disproportionate hardships in the pandemic.

REFERENCES

1. World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard. https://covid19.who.int/. Accessed June 23, 2021.
2. World Health Organization. Impact of COVID-19 on people's livelihoods, their health and our food systems: joint statement by ILO, FAO, IFAD and WHO. https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people%27s-livelihoods-their-health-and-our-food-systems. Accessed June 23, 2021.
3. Peckham H, de Gruijter NM, Raine C, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. *Nat Commun*. 2020;11:6317.
4. Islam N, Shkolnikov VM, Acosta RJ, et al. Excess deaths associated with covid-19 pandemic in 2020: age and sex disaggregated time series analysis in 29 high income countries. *BMJ*. 2021;373:n1137.
5. Connor J, Madhavan S, Mokashi M, et al. Health risks and outcomes that disproportionately affect women during the Covid-19 pandemic: a review. *Soc Sci Med.* 2020;266:113354.
6. Ahmed SB, Dumanski SM. Sex, gender and COVID-19: a call to action. *Can J Public Health*. 2020;111:980-983.
7. World Health Organization. Gender and COVID-19: advocacy brief – 14 May 2020. https://www.who.int/publications/i/item/WHO-2019-nCoV-Advocacy_brief-Gender-2020.1. Accessed June 23, 2021.
8. United Nations. Policy brief: the impact of COVID-19 on women – 9 April 2020. https://unsdg.un.org/resources/policy-brief-impact-covid-19-women. Accessed June 23, 2021.
9. UN Women 2020. From insight to action: gender inequality in the wake of COVID-19. https://www.unwomen.org/en/digital-library/publications/2020/09/gender-equality-in-the-wake-of-covid-19. Accessed June 23, 2021.

10. World Bank Group. Gender dimensions of the COVID-19 pandemic. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/618731587147227244/gender-dimensions-of-the-covid-19-pandemic. Accessed June 23, 2021.
11. United States Census Bureau. Custodial Mothers and Fathers and Their Child Support:2015. https://www.census.gov/library/publications/2018/demo/p60-262.html. Accessed June 23, 2021.
12. Piquero AR, Jennings WG, Jemison E, et al. Domestic violence during the COVID-19 pandemic - Evidence from a systematic review and meta-analysis. *J Crim Justice*. 2021;74:101806.
13. Flentje A, Obedin-Maliver J, Lubensky ME, et al. Depression and anxiety changes among sexual and gender minority people coinciding with onset of COVID-19 pandemic. *J Gen Intern Med* 2020;35:2788-90.
14. The Lives and livelihoods of many in the LGBTQ community are at risk amidst the COVID-19 crisis. Human Rights Campaign Foundation; 2020. https://assets2.hrc.org/files/assets/resources/COVID19-IssueBrief-032020-FINAL.pdf. Accessed June 23, 2021.
15. Lindau ST, Makelarski JA, Boyd K, et al. Change in health-related socioeconomic risk factors and mental health during the early phase of the COVID-19 pandemic: a national survey of U.S. women. *J Womens Health*. 2021;30:502-513.
16. García-Fernández L, Romero-Ferreiro V, Padilla S, et al. Gender differences in emotional response to the COVID-19 outbreak in Spain. *Brain Behav*. 2021;11:e01934.
17. Statistics Canada. Gender differences in mental health during the COVID-19 pandemic. https://www150.statcan.gc.ca/n1/pub/45-28-0001/2020001/article/00047-eng.htm. Accessed June 23, 2021.

18. Centre for Addiction and Mental Health. COVID-19 pandemic adversely affecting mental health of women and people with children. https://www.camh.ca/en/camh-news-and-stories/covid-19-pandemic-adversely-affecting-mental-health-of-women-and-people-with-children. Accessed June 23, 2021.
19. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry.* 1994;51:8-19.
20. Salk RH, Hyde JS, Abramson LY. Gender differences in depression in representative national samples: meta-analysis of diagnosis and symptoms. *Psychol Bull*. 2017;143:783-822.
21. Patil PA, Porche MV, Shippen NA, et al. Which girls, which boys? The intersectional risk for depression by race and ethnicity, and gender in the U.S. *Clin Psychol Rev.* 2018;66:51-68.
22. McLean CP, Anderson ER. Brave men and timid women? A review of the gender differences in fear and anxiety. *Clin Psychol Rev.* 2009;29:496-505.
23. McLean CP, Asnaani A, Litz BT, et al. Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. *J Psychiatr Res.* 2011;45:1027-1035.
24. Thombs BD, Bonardi O, Rice DB, et al. Curating evidence on mental health during COVID-19: a living systematic review. *J Psychosom Res* 2020;133;110113.
25. Living systematic review of mental health in COVID-19. https://www.depressd.ca/covid-19-mental-health**.** Accessed June 23, 2021.
26. Sun Y, Wu Y, Bonardi O, et al. Comparison of mental health symptoms prior to and during COVID-19: evidence from a living systematic review and meta-analysis. *MedRxiv*. https://doi.org/10.1101/2021.05.10.21256920.
27. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71.

28. World Health Organization. Rolling updates on coronavirus disease (COVID-19) 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen. Accessed June 23, 2021.
29. World Health Organization. Gender and health. https://www.who.int/health-topics/gender#tab=tab_1. Accessed June 23, 2021.
30. Joanna Briggs Institute. The Joanna Briggs Institute Critical Appraisal tools for use in JBI Systematic Reviews: Checklist for Prevalence Studies. https://jbi.global/critical-appraisal-tools. Accessed June 23, 2021.
31. Hedges LV. Estimation of effect from a series of independent experiments. *Pyschol Bull*. 1982;92:490-499.
32. Borenstein M, Hedges LV, Higgins JPT, et al. Chapter 4: Effect sizes based on means. In Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (editors). Introduction to meta-analysis. Wiley & Sons: West Sussex, UK. 2009.
33. Agresti A, Coull BA. Approximate is better than “exact” for interval estimation of binomial proportions. *Am Stat* 1998;52:119-26.
34. Newcombe RG. Improved confidence intervals for the difference between binomial proportions based on paired data. *Stat Med* 1998;17:2635-2650.
35. Newcombe RG. Estimating the difference between differences: measurement of additive scale interaction for proportions. *Stat Med.* 2001;20:2885-2893.
36. Schwarzer G. meta: An R package for meta-analysis. *R News*. 2007;7:40-45.
37. Pierce M, Hope H, Ford T, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. *Lancet Psychiatry.* 2020;7:883-892.
38. Daly M, Robinson E. Longitudinal changes in psychological distress in the UK from 2019 to September 2020 during the COVID-19 pandemic: Evidence from a large nationally representative study. *Psychiatry Res.* 2021;300:113920.

39. van der Velden PG, Contino C, Das M, et al. Anxiety and depression symptoms, and lack of emotional support among the general population before and during the COVID-19 pandemic. a prospective national study on prevalence and risk factors. *J Affect Disord.* 2020;277:540-548.

40. Van der Velden PG, Marchand M, Das M, et al. The prevalence, incidence and risk factors of mental health problems and mental health services use before and 9 months after the COVID-19 outbreak among the general Dutch population. A 3-wave prospective study. *MedRixv.* 2021. https://doi.org/10.1101/2021.02.27.21251952\.

41. Wang ZH, Qi SG, Zhang H. 新型冠状病毒肺炎对社区老年人焦虑症状的影响 [Impact of the COVID-19 epidemic on anxiety among the elderly in community]. *Natl Med J China.* 2020;100:3179-85.

42. Rimfeld K, Malanchini M, Allegrini AG, et al. Genetic correlates of psychological responses to the COVID-19 crisis in young adult twins in Great Britain. *Behav Genet.* 2021;51(2):110-124.

43. Shanahan L, Steinhoff A, Bechtiger L, et al. Emotional distress in young adults during the COVID-19 pandemic: evidence of risk and resilience from a longitudinal cohort study. *Psychol Med.* 2020;1-10.

44. Magson NR, Freeman JUA, Rapee RM, et al. Risk and protective factors for prospective changes in adolescent mental health during the COVID-19 pandemic. *J Youth Adolesc.* 2021;50:44-57.

45. Dong XL. 新型冠状病毒肺炎疫情对师范类大学生心理健康的影响研究 [Influence study of COVID-2019 on mental health of normal college students]. *Xin Li Yue Kan*. 2020;20:15.

46. Saraswathi I, Saikarthik J, Kumar KS, et al. Impact of COVID-19 outbreak on the mental health status of undergraduate medical students in a COVID-19 treating medical college: a prospective longitudinal study. *PeerJ.* 2020;8:e10164.

47. Savage MJ, James R, Magistro D, et al. Mental health and movement behaviour during the COVID-19 pandemic in UK university students: Prospective cohort study. *Ment Health Phy Act*. 2020;19:100357.

48. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41:582-592.

49. Riecher-Rössler A. Prospects for the classification of mental disorders in women. *Eur Psychiatry.* 2010;25:189-196.

50. Boyd A, van de Velde S, Vilagut G, et al. Gender differences in mental disorders and suicidality in Europe: results from a large cross-sectional population-based study. *J Affect Disord*. 2015;173:245-254.

51. Seedat S, Scott KM, Angermeyer MC, et al. Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. *Arch Gen Psychiatry*. 2009;66:785-795.

52. Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol*. 2011;21:655-679.

53. Kuehner C. Why is depression more common among women than men? *Lancet Psychiatry*. 2016;4:146-158.

54. Riecher-Rössler A. Sex and gender differences in mental disorders. *Lancet Psychiatry*. 2016;4:8-9.

55. Bentall R. Has the pandemic really cause a ‘tsunami’ of mental health problems? *The Guardian*. February 9, 2021. https://www.theguardian.com/commentisfree/2021/feb/09/pandemic-mental-health-problems-research-coronavirus. Accessed June 23, 2021.

56. Marmet S, Wicki M, Gmel G, et al. The psychological impact of the COVID-19 crisis on young Swiss men participating in a cohort study. *PsyArXiv* 2020. Preprint. https://doi.org/10.31234/osf.io/kwxhd.

FIGURE LEGENDS

Figure 1. PRISMA flow diagram.

Figures 2a-2b. Forest plots of standardized mean difference of the difference in change in continuous anxiety symptom scores (2a) and the difference in change in proportion above a cut-off (2b) between females or women and males or men. Positive numbers indicate more negative change in anxiety symptoms in females or women compared to males or men.

Figures 3a-3b. Forest plots of standardized mean difference of the difference in change in continuous depression symptom scores (3a) and the difference in change in proportion above a cut-off (3b) between females or women and males or men. Positive numbers indicate more negative change in depression symptoms in females or women compared to males or men.

Figures 4a-4c. Forest plots of standardized mean difference of the difference in change in continuous general mental health scores (4a) and the difference in change in proportion above a cut-off (4b) between females or women and males or men. Positive numbers indicate more negative change in general mental health in females or women compared to males or men. Figures 4a and 4b reflect COVID-19 mental health measured in early 2020, whereas 4c reflects measurements from late 2020 for Daly[38] and van der Velden.[39]

Figures 5a-5b. Forest plots of standardized mean difference of the difference in change in continuous stress scores (5a) and the difference in change in proportion above a cut-off (5b) between females or women and males or men. Positive numbers indicate more negative change in stress in females or women compared to males or men.

Figure 6. Illustration of the magnitude of change for SMD = 0.15 assuming a normal distribution. The hypothetical blue distribution represents pre-COVID-19 scores, and the grey distribution represents post-COVID-19 scores with a mean symptom increase of SMD = 0.15.

Appendix 1. Search Strategies

Search strategies can be found in the project folder on the Open Science Framework:

https://osf.io/2zh9n/

Appendix 2. Title and Abstract and Full-Text Review Inclusion and Exclusion Criteria Coding Guides

Title and Abstract Coding Criteria

MENTAL HEALTH SYMPTOM CHANGES CODING CRITERIA:

No: not original human data or a case study or case series. If it is clear from the title and abstract that the article is not an original report of primary data, but, for example, a letter, editorial, systematic review or meta-analysis, or it is a single case study or case series, then it is excluded. Studies reporting only on animal, cellular, or genetic data are also excluded. Conference abstracts are included.

No: not a study of any population affected by the COVID-19 outbreak. If it is clear from the title or abstract that the study is not about any population affected by the COVID-19 outbreak, it is excluded. Studies that include fewer than 100 subjects, are excluded.

No: not a study which reports mental health symptom changes over a demarcated period. If it is clear from the title or abstract that the study does not report proportions of participants meeting diagnostic criteria using a validated diagnostic interview or validated mental health scale, or proportions of symptoms (based on a threshold or measured continuously) prior to and after a delineated event related to COVID-19,then it will be excluded. Delineated events for pre-post comparisons may include pandemic announcements, social isolation regulations, etc. Pre- and post- samples must include the same cohort, not representative samples. Pre- and post- samples should have less than 10% missing data, or should statistically account for missing data if N between samples varies by more than 10%.

Yes: study eligible for inclusion in full-text review.

Full Text Coding Criteria

MENTAL HEALTH SYMPTOM CHANGES CODING CRITERIA:

No: not original human data or a case study or case series. If the article is not an original report of primary data, but, for example, a letter, editorial, systematic review or meta-analysis, or it is a single case study or case series, then it is excluded. Studies reporting only on animal, cellular, or genetic data are also excluded. Conference abstracts are included.

No: not a study of any population affected by the COVID-19 outbreak. If the study is not about any population affected by the COVID-19 outbreak, it is excluded. Studies that include fewer than 100 subjects, are excluded.

No: not a study which reports mental health symptom changes over a demarcated period. If it the study does not report proportions of participants meeting diagnostic criteria using a validated diagnostic interview or validated mental health scale, or proportions of symptoms (based on a threshold or measured continuously) prior to and after a delineated event related to COVID-19,then it will be excluded. Delineated events for pre-post comparisons may include pandemic announcements, social isolation regulations, etc. Delineated events for pre-post comparisons may include pandemic announcements, social isolation regulations, etc. Pre- and post- samples must include the same cohort, not representative samples. Pre- and post- samples should have less than 10% missing data, or should statistically account for missing data if N between samples varies by more than 10%.

Yes: study eligible for inclusion in systematic review.

Appendix 3: Adequacy of Study Methods and Reporting

Q1. Was the sample frame appropriate to address the target population?

Yes: The sampling frame was a true or close representation of the target population.
No: The sampling frame was NOT a true or close representation of the target population.
Unclear: Not enough information provided to determine.

Q2. Were study participants recruited in an appropriate way?

Yes: A census was undertaken, OR, some form of random selection was used to select the sample (e.g. simple random sampling, stratified random sampling, cluster sampling, systematic sampling).
No: A census was NOT undertaken, AND some form of random selection was NOT used to select the sample.
Unclear: Not enough information provided to determine.

Q3. Was the sample size adequate?

Yes: There is evidence that the authors conducted a sample size calculation to determine an adequate sample size OR the study was large enough (e.g., a large national survey) whereby a sample size calculation is not required. In these cases, sample size can be considered adequate. If at least 200 participants are included for continuous outcomes and 250 for proportions, this is considered low risk.
No: The authors did not reach their intended sample size, or no sample size calculation is provided and there are < 100 participants for continuous outcomes, or < 125 for proportions.
Unclear: No sample size calculation is provided, and between 100-199 participants are included for continuous outcomes or between 125-249 for proportions.

Q4. Were the study participants and setting described in detail?

Yes: Data included age, sex, and at least 1 socioeconomic indicator (e.g., income, education, work status).
No: The minimum sociodemographic variables have not been reported.
Unclear: Not stated

Q5. Was the response rate adequate and was the data analysis conducted with sufficient coverage?

Yes: The overall response rate or response rate for intended subgroups was >/=75%, OR, an analysis was performed that established that there was not a substantive difference in relevant demographic characteristics between responders and non-responders within a subgroup (if non-response too high (e.g., > 50%), code “No”)
No: The overall response rate or response rate for subgroups was <75%, and if any analysis comparing responders and non-responders was done, it showed a meaningful difference in relevant demographic characteristics between responders and non-responders.
Unclear: Not enough information provided to determine.

Q6. Were valid methods used for the identification of the outcome variable?

Yes: The study instrument had been shown to have reliability and validity, e.g., test-retest, piloting, validation in a previous study, etc.
No: The study instrument had NOT been shown to have reliability or validity.
Unclear: Not stated.

Q7. Was the mental health outcome measured in a standard, reliable way for all participants?

Yes: All self-report data were collected directly from the participants. Any clinical interview data includes at least information about the interviewers' level of education or training received. The same mode of data collection was used for all participants. All aspects of this question must be present (where relevant).
No: In some instances, data were collected from a proxy (e.g., a spouse). The qualifications of clinical interviewers are not reported or not appropriate. The same mode of data collection was NOT used for all participants. If any aspects of this item are absent, it is high risk.
Unclear: Not stated.

Q8. Was there appropriate statistical analysis?

Yes: Continuous variables report (1) mean (SD) of change or (2) pre mean (SD) and post mean (SD) with/out correlation between pre and post scores. For dichotomous variables, numerator, denominator, and percentages are clearly reported. Continuous variables are not artificially dichotomized. The statistical analyses section is detailed enough for readers to understand change scores (see STROBE reporting guidelines, if necessary).
No: Continuous variables do not include a report of the (1) mean (SD) of change or (2) pre mean (SD) and post mean (SD) with/out correlation between pre and post scores. For dichotomous variables, the numerator, denominator, or percentages are not clearly reported. The statistical analyses section does not clearly describe the methods used to assess change scores.

Q9. Was the follow-up rate adequate, and if not, was the low follow-up rate managed appropriately?

Yes: At least 75% of those who participated in the pre-COVID-19 assessment(s) provided follow-up responses and had their responses included in the follow-up, OR, an analysis was performed that showed no substantive difference in relevant demographic characteristics between participants who stayed in the study and drop-outs (if dropout too high (e.g. > 50%), code "No").
No: Less than 75% of those participated in the pre-COVID-19 assessment(s) provided responses and had their responses included in the follow-up, and if any analysis comparing participants who stayed in the study and drop-outs was done, it showed a substantive difference in relevant demographic characteristics between the two groups.
Unclear: Not stated.

Appendix 4. Confidence Intervals for Proportion Change When Using 30% and 70% of pre-COVID-19 Cases were Cases during COVID-19: Calculated for Studies that Did Not Provide 95% Confidence Intervals for Change

First Author	Outcome Domain	Sex/Gender	95% CI with 30%	95% CI with 70%
Dong[45]	General Mental Health	Females/Women	-0.10, -0.06	-0.10, -0.07
		Males/Men	-0.10, -0.04	-0.10, -0.04
Saraswathi[46]	Anxiety Symptoms	Females/Women	-0.24, -0.03	-0.22, -0.06
		Males/Men	-0.23, 0.06	-0.20, 0.02
	Depression Symptoms	Females/Women	-0.10, 0.14	-0.05, 0.09
		Males/Men	-0.25, 0.05	-0.21, 0.01
	Stress	Females/Women	-0.12, 0.06	-0.09, 0.04
		Males/Men	-0.20, 0.08	-0.16, 0.03
van der Velden[39]	General Mental Health	Females/Women	-0.02, 0.03	-0.01, 0.02
		Males/Men	-0.03, 0.01	-0.02, 0.00
van der Velden[40]	General Mental Health	Females/Women	-0.01, 0.03	-0.00, 0.03
		Males/Men	-0.03, 0.01	-0.03, 0.00
Wang[41]	Anxiety Symptoms	Females/Women	-0.06, -0.04	-0.06, -0.04
		Males/Men	-0.06, -0.04	-0.06, -0.04

Table 1. Characteristics of included studies (N=11)

First Author	Outcome Domains				Description of Participants	Country	Pre- and Post-COVID-19 Dates of Data Collection	N Females or Women (F/W) and Males or Men (M/M)	Mean (SD) Participant Age or Age Range (%)	Use of Sex or Gender
	Anxiety Symptoms	Depression Symptoms	General Mental Health	Stress						
Dong[45]			SCL-90-R		First-year undergraduate students from a single university recruited online	China	09/2019 NR/2020	F/W: 3,162 M/M: 923	19 (1)	Gender
Magson[44]	SCAS	SMFQ			Adolescents aged 13-16 years who enrolled in a longitudinal cohort 4 years prior	Australia	NR/2019 05/2020	Girls: 126 Boys: 122	14 (1)	Inconsistent
Pierce[37] Daly[38]			GHQ-12		National probability-based sample of adults aged ≥ 18 years (United Kingdom Household Longitudinal Study)	United Kingdom	Pre-COVID-19 waves[a] 04-09/2020	F/W: 7,181[b] M/M: 8,195[b]	18-34 (12)[c] 35-49 (22)[c] 50-64 (34)[c]	Gender
								F/W: 6,380 M/M: 4,538	65+ (32)[c]	Inconsistent
Rimfeld[42]	GAD-7	SMFQ			Adult twins born between 1994-1996 who were enrolled in a longitudinal cohort at age 18 months	United Kingdom	NR/2018 04-05/2020	F/W: 2,513-2,578 M/M: 1,050-1,116	24-26 (100%)	Inconsistent
Saraswathi[46]	DASS-21 Anxiety	DASS-21 Depression		DASS-21 Stress	Convenience sample of undergraduate university medical students	India	12/2019 06/2020	F/W: 139 M/M: 78	20 (2)	Inconsistent
Savage[47]			WEMWBS	PSS	Undergraduate students from single university recruited by email invitation and enrolled in an ongoing longitudinal study	United Kingdom	10/2019 04/2020	F/W: 154 M/M: 60	18-21 (64) 22-25 (22) 26-35 (8) 35+ (6)	Inconsistent
Shanahan[43]				PSS	Young adults who participated in a longitudinal cohort of the Zurich Project on the Social Development from Childhood to Adulthood	Switzerland	NR/2018 04/2020	F/W: 378 M/M: 408	22 (0)	Sex
van der Velden[39] van der Velden[40]			MHI-5		National probability-based sample of adults aged ≥ 18 years (Longitudinal Internet Studies for the Social Sciences)	The Netherlands	03/2019 11-12/2019	F/W: 2,020 M/M: 1,963	18-34 (25)[d] 35-49 (23)[d] 50-64 (26)[d]	Gender
							03/2020 11-12/2020	F/W: 2, 062 M/M: 2,002	65+(26)[d]	Sex
Wang[41]	GAD-2				Volunteers recruited via publicity from a nationally representative sample of adults aged ≥ 65 years who had completed pre-COVID-19 measures	China	10/2019 05/2020	F/W: 3,599 M/M: 2,868	65-69 (45) 70-74 (29) 75-79 (15) >80 (12)	Gender

DASS-21 Anxiety = Depression, Anxiety, and Stress Scale - Anxiety subscale; DASS-21 Depression = Depression, Anxiety, and Stress Scale - Depression subscale; DASS-21 Stress = Depression, Anxiety, and Stress Scale - Stress subscale; GAD-2 = Generalized Anxiety Disorder-2; GAD-7 = Generalized Anxiety Disorder-7; GHQ-12 = General Health Questionnaire-12; MHI-5 = Mental Health Index-5; PSS = Perceived Stress Scale; SCAS = Spence Children's Anxiety Scale; SCL-90-R = Symptom Check List-90 Revised; SMFQ = Short Mood and Feelings Questionnaire; WEMWBS = Warwick Edinburgh Mental Wellbeing Scale.

[a]Analyses compared COVID-19 symptom levels to preceding trends across multiple assessments. [b]Number included in fixed effects regression analysis from where the majority of data were extracted. [c]Age groups reported for Daly[38]; for Pierce,[37] 16-24 = 9%, 25-34 = 11%, 35-44 = 16%, 45-54 = 20%, 55-69 = 29%, 70+ =15%.[d]Based on van der Velden.[39]

Table 2. Adequacy of methods and reporting of included studies (N=11)

Author	Appropriate sample frame	Appropriate participant recruitment	Adequate sample size	Subjects and setting adequately described	Adequate response rate and data analysis with sufficient coverage	Valid methods for identification of outcome variable	Standard, reliable outcome measurement	Appropriate statistical analysis	Adequate follow-up response rate/ appropriate management of low response rate
Dong[45]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Magson[44]	No	Unclear	Yes	Yes	Unclear	Yes	Yes	Yes	No
Pierce[37] Daly[38]	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
Rimfeld[42]	Yes	Unclear	Yes	Yes	Unclear	Yes	Yes	Yes	Unclear
Saraswathi[46]	No	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes
Savage[47]	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No
Shanahan[43]	No	Unclear	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
van der Velden[39] van der Velden[40]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wang[41]	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	No

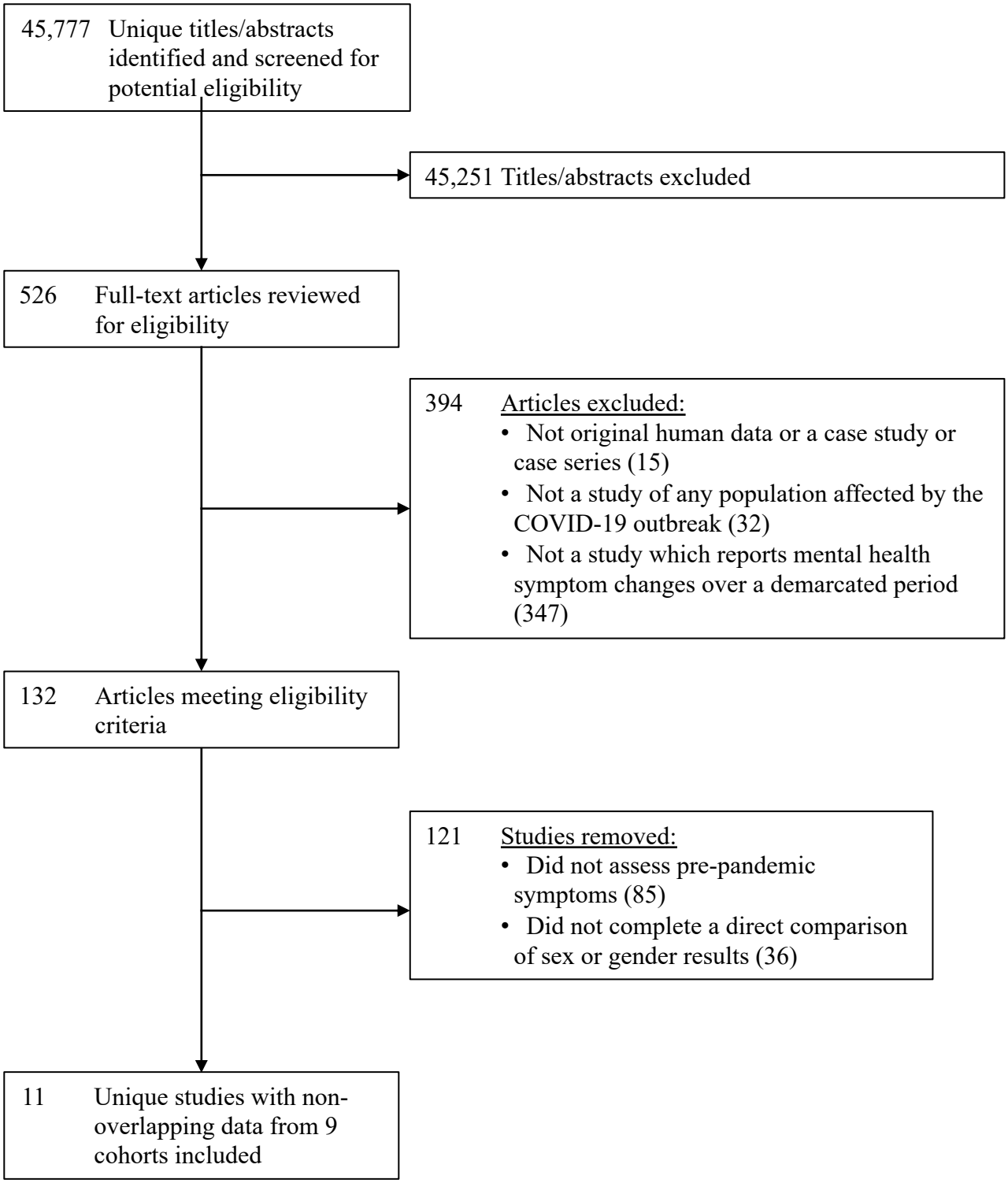
Table 3. Outcomes from Included Studies by Sex or Gender[a]

First Author	Pre- and Post-COVID-19 Data Collection	Sex or Gender	N	Continuous Outcome Measure	Pre-COVID-19 Mean (SD)	Post-COVID-19 Mean (SD)	Mean (SD) Change	Hedges' g Standardized Mean Difference (95% CI)	Dichotomous Outcome Measure	% Pre-COVID-19 (95% CI)	% Post-COVID-19 (95% CI)	% Change (95%CI)
Anxiety Symptoms												
Magson[44]	NR/2019 05/2020	Females/Women	126	SCAS-C	5.55 (4.05)	6.52 (4.31)	0.97 (4.18)	0.23 (-0.02, 0.48)	----------	----------	----------	----------
		Males/Men	122		3.63 (3.13)	3.64 (3.16)	0.01 (3.14)	0.00 (-0.25, 0.25)				
Rimfeld[42]	NR/2018 04-05/2020	Females/Women	2,513	GAD-7	8.15 (7.53)	9.69 (7.69)	1.54 (7.61)	0.20 (0.15, 0.26)	----------	----------	----------	----------
		Males/Men	1,050		5.88 (6.66)	6.30 (6.58)	0.42 (6.62)	0.06 (-0.02, 0.15)				
Saraswathi[46]	12/2019 06/2020	Females/Women	139	DASS-21 Anxiety	4.59 (6.29)	5.94 (6.93)	1.35 (6.62)	0.20 (-0.03, 0.44)	DASS-21 Anxiety > 7	18.7 (13.1, 26.0)	32.4 (25.2, 40.5)	13.7 (4.4, 22.7)
		Males/Men	78		4.62 (6.04)	6.41 (7.50)	1.79 (6.81)	0.26 (-0.05, 0.57)		25.6 (17.3, 36.3)	34.6 (25.0, 45.7)	9.0 (-4.0, 21.5)
Wang[41]	10/2019 05/2020	Females/ Women	3,599	----------	----------	----------	----------	----------	GAD-2 ≥ 2	5.6 (4.8, 6.3)	10.7 (9.7, 11.7)	5.1 (4.1, 6.2)
		Males/Men	2,868							4.2 (3.5, 4.9)	9.4 (8.3, 10.4)	5.2 (4.1, 6.3)
Depression Symptoms												
Magson[44]	NR/2019 05/2020	Females/Women	126	SMFQ-C	4.77 (5.00)	8.16 (6.46)	3.39 (5.78)	0.58 (0.33, 0.83)	----------	----------	----------	----------
		Males/Men	122		2.81 (3.18)	4.02 (4.76)	1.21 (4.05)	0.30 (0.05, 0.55)				
Rimfeld[42]	NR/2018 04-05/2020	Females/Women	2,578	SMFQ	4.65 (4.20)	4.81 (4.07)	0.16 (4.14)	0.04 (-0.02, 0.09)	----------	----------	----------	----------
		Males/Men	1,116		3.71 (3.70)	3.33 (3.40)	-0.38 (3.55)	-0.11 (-0.19, -0.02)				
Saraswathi[46]	12/2019 06/2020	Females/Women	139	DASS-21 Depression	7.71 (7.57)	7.94 (8.77)	0.23 (8.19)	0.03 (-0.21, 0.26)	DASS-21 Depression > 9	36.7 (29.1, 45.0)	34.5 (27.1, 42.8)	-2.2 (-11.7, 7.4)
		Males/Men	78		7.28 (8.40)	8.54 (9.17)	1.26 (8.79)	0.14 (-0.17, 0.45)		26.9 (18.3, 37.7)	37.2 (27.3, 48.3)	10.3 (-2.9, 22.9)
General Mental Health												
Dong[45]	09/2019 NR/2020	Females/Women	3,162-3,277	----------	----------	----------	----------	----------	SCL-90-R ≥ 160	19.7 (18.4, 21.1)	27.9 (26.4, 29.5)	8.2 (6.3, 10.0)
		Males/Men	923-1,064							14.3 (12.3, 16.5)	21.2 (18.7, 24.0)	6.9 (4.0, 9.9)
Pierce[37]	Pre-COVID-19 Waves	Females/Women	7,181[2,b]	GHQ-12	12.00 (5.91)	13.60 (7.14)	1.60 (6.55)[c]	0.24 (0.21, 0.28)	GHQ-12 ≥ 4	24.5 (22.5, 26.4)[e]	36.8 (34.8, 38.9)[e]	12.4 (9.9, 14.9)[e]
Daly[38]			6,380				0.88 (NR)[d]	0.13 (0.10, 0.17)		24.5 (22.5, 26.4)[e]	25.0 (23.3, 26.8)[e]	0.5 (-1.8, 2.9)[e]

	04/2020 09/2020	Males/Men	8,195[2] [2] 4,538		10.80 (4.99)	11.50 (5.75)	0.70 (5.38)[c] 0.03 (NR)[d]	0.13 (0.10, 0.16) 0.01 (-0.03, 0.04)		16.7 (14.6, 18.7)[e] 16.7 (14.6, 18.7)[e]	21.1 (19.0, 23.3)[e] 16.0 (14.0, 17.9)[e]	4.5 (2.0, 7.0)[e] -0.7 (-2.9, 1.5)[e]
Savage[47]	10/2019 04/2020	Females/Women	154	WEMWBS[f]	43.00 (9.00)	40.00 (10.00)	-3.00 (9.51)	0.31 (0.09, 0.54)	----------	----------	----------	----------
		Males/Men	60		47.00 (9.00)	44.00 (10.00)	-3.00 (9.51)	0.31 (-0.05, 0.67)				
van der Velden[39]	03/2019 11-12/2019	Females/Women	2,020 2,062	----------	----------	----------	----------	----------	MHI-5 ≤ 59	18.9 (17.3, 20.7) 19.1 (17.4, 20.8)	18.3 (16.7, 20.1) 17.8 (16.2, 19.5)	-0.6 (-2.5, 1.3) -1.3 (-3.1, 0.6)
van der Velden[40]	03/2020 11-12/2020	Males/Men	1,962-1,963 2,002							14.6 (13.1, 16.3) 14.7 (13.2, 16.3)	15.6 (14.1, 17.3) 15.9 (14.4, 17.6)	1.0 (-0.8, 2.7) 1.2 (-0.5, 3.0)
Stress												
Saraswathi[46]	12/2019 06/2020	Females/Women	139	DASS-21 Stress	6.95 (7.22)	8.88 (7.99)	1.93 (7.61)	0.25 (0.02, 0.49)	DASS-21 Stress > 14	19.4 (13.7, 26.8)	22.3 (16.2, 29.9)	2.9 (-4.9, 10.6)
		Males/Men	78		7.95 (7.54)	10.08 (8.50)	2.13 (8.03)	0.26 (-0.05, 0.58)		23.1 (15.1, 33.6)	29.5 (20.5, 40.4)	6.4 (-5.6, 18.2)
Savage[47]	10/2019 04/2020	Females/Women	154	PSS	21.00 (7.00)	24.00 (7.00)	3.00 (7.00)	0.43 (0.20, 0.65)	----------	----------	----------	----------
		Males/Men	60		17.00 (6.00)	21.00 (7.00)	4.00 (6.52)	0.61 (0.24, 0.97)				
Shanahan[43]	NR/2018 04/2020	Females/Women	378	PSS	3.02 (0.98)	3.10 (0.94)	0.08 (0.96)	0.08 (-0.06, 0.23)	----------	----------	----------	----------
		Males/Men	408		2.57 (0.86)	2.74 (0.86)	0.17 (0.86)	0.20 (0.06, 0.34)				

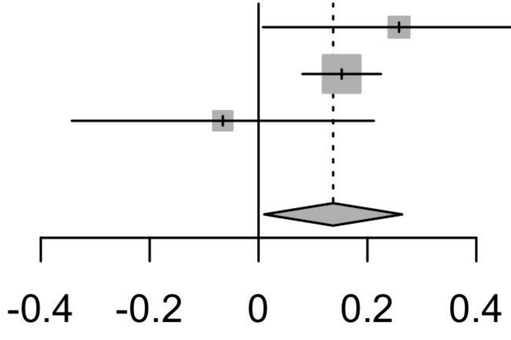
Stress Scale - Stress subscale; GAD-2 = Generalized Anxiety Disorder-2; GAD-7 = Generalized Anxiety Disorder-7; GHQ-12 = General Health Questionnaire-12; MHI-5 = Mental Health Index-5; PSS = Perceived Stress Scale; SCAS = Spence Children's Anxiety Scale; SCL-90-R = Symptom Check List-90 Revised; SMFQ = Short Mood and Feelings Questionnaire; WEMWBS = Warwick Edinburgh Mental Wellbeing Scale.

[a]Positive Hedge's g sizes and increases in proportions above a threshold indicate worse mental health in COVID-19 compared to pre-COVID-19. Effects for measures where high scores = positive outcomes were reversed to reflect this. [b] Number included in fixed effects regression analysis from where majority of data were extracted. [c]Based on difference between 2020 and 2019 outcomes. [d]Based on estimate from fixes effects regression model that estimates within-person change accounting for pre-COVID-19 trends.[e]Included proportion outcomes from Daly[38] since they reported for two time points. [f] Higher scale scores reflect better mental health; thus, direction of effect sizes reversed.


45,777 Unique titles/abstracts identified and screened for potential eligibility
45,251 Titles/abstracts excluded
526 Full-text articles reviewed for eligibility
394 Articles excluded:
• Not original human data or a case study or case series (15)
• Not a study of any population affected by the COVID-19 outbreak (32)
• Not a study which reports mental health symptom changes over a demarcated period (347)
132 Articles meeting eligibility criteria
121 Studies removed:
• Did not assess pre-pandemic symptoms (85)
• Did not complete a direct comparison of sex or gender results (36)
11 Unique studies with non-overlapping data from 9 cohorts included

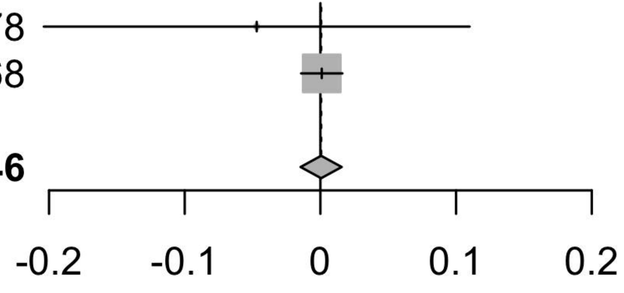
Study	Females/Women Total	Mean	SD	Males/Men Total	Mean	SD	SMD	95% CI
Magson, 2021	126	0.97	4.18	122	0.01	3.14	0.26	[0.01; 0.51]
Rimfeld, 2021	2513	1.54	7.61	1050	0.42	6.62	0.15	[0.08; 0.22]
Saraswathi, 2020	139	1.35	6.62	78	1.79	6.81	-0.07	[-0.34; 0.21]
Random effects model	**2778**			**1250**			**0.14**	**[0.01; 0.26]**

Heterogeneity: I^2 = 34%



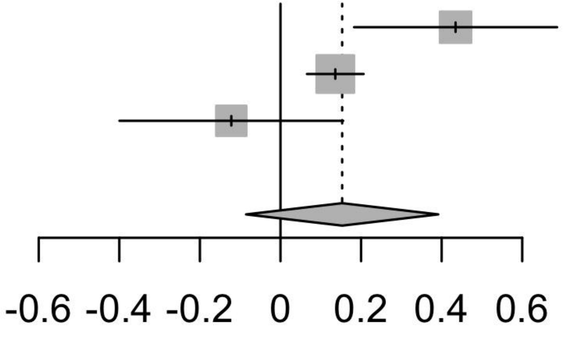
Study	Females/Women N	Males/Men N	Difference in Changes	Difference	95% CI
Saraswathi, 2020	139	78		-0.05	[-0.20; 0.11]
Wang, 2020	3599	2868		0.00	[-0.01; 0.02]
Random effects model	**3738**	**2946**		**0.00**	**[-0.01; 0.02]**

Heterogeneity: I^2 = 0%

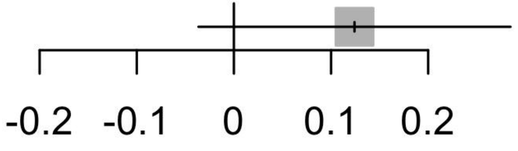


Study	Females/Women Total	Mean	SD	Males/Men Total	Mean	SD	SMD	95% CI
Magson, 2021	126	3.39	5.78	122	1.21	4.05	0.43	[0.18; 0.69]
Rimfeld, 2021	2578	0.16	4.14	1116	-0.38	3.55	0.14	[0.07; 0.21]
Saraswathi, 2020	139	0.23	8.19	78	1.26	8.79	-0.12	[-0.40; 0.16]
Random effects model	**2843**			**1316**			**0.15**	**[-0.09; 0.39]**

Heterogeneity: I^2 = 77%

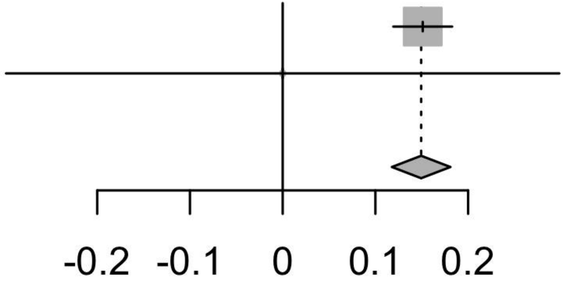


Study	Females/Women N	Males/Men N	Difference in Changes	Difference	95% CI
Saraswathi, 2020	139	78		0.12	[-0.04; 0.28]



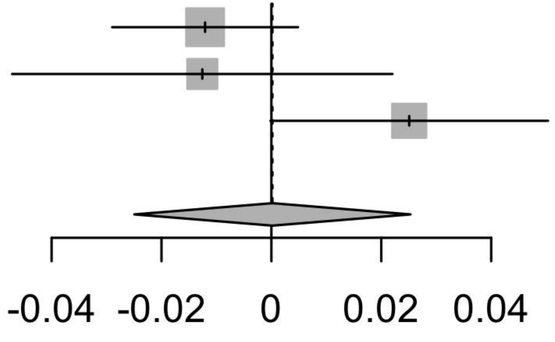
Study	Females/Women Total	Mean	SD	Males/Men Total	Mean	SD	SMD	95% CI
Pierce, 2020	7181	1.60	6.55	8195	0.70	5.38	0.15	[0.12; 0.18]
Savage, 2020	154	-3.00	9.51	60	-3.00	9.51	0.00	[-0.30; 0.30]
Random effects model	**7335**			**8255**			**0.15**	**[0.12; 0.18]**

Heterogeneity: I^2 = 0%



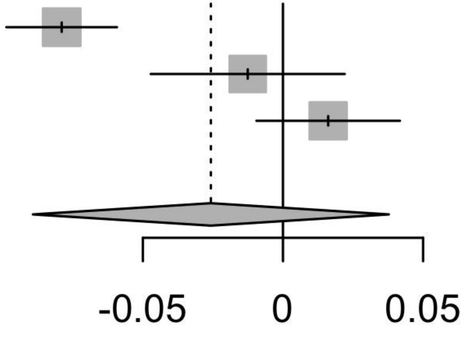
Study	Females/Women N	Males/Men N	Difference in Changes	Difference	95% CI
Daly, 2021	6380	4538		-0.01	[-0.03; 0.00]
Dong, 2020	3162	923		-0.01	[-0.05; 0.02]
van der Velden, 2021	2062	2002		0.03	[0.00; 0.05]
Random effects model	**11604**	**7463**		**0.00**	**[-0.02; 0.03]**

Heterogeneity: I^2 = 67%



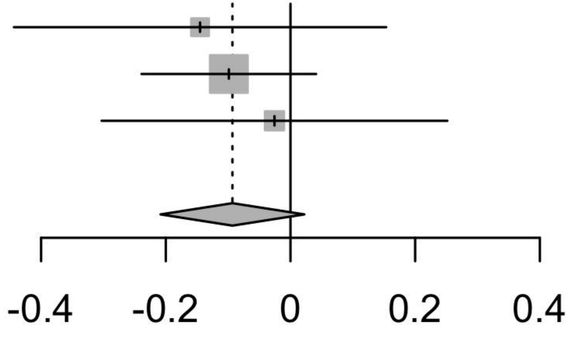
Study	Females/Women N	Males/Men N	Difference in Changes	Difference	95% CI
Daly, 2021	6380	4538		-0.08	[-0.10; -0.06]
Dong, 2020	3162	923		-0.01	[-0.05; 0.02]
van der Velden, 2020	2020	1962		0.02	[-0.01; 0.04]
Random effects model	**11562**	**7423**		**-0.03**	**[-0.09; 0.04]**

Heterogeneity: I^2 = 94%

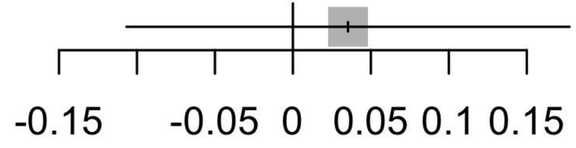


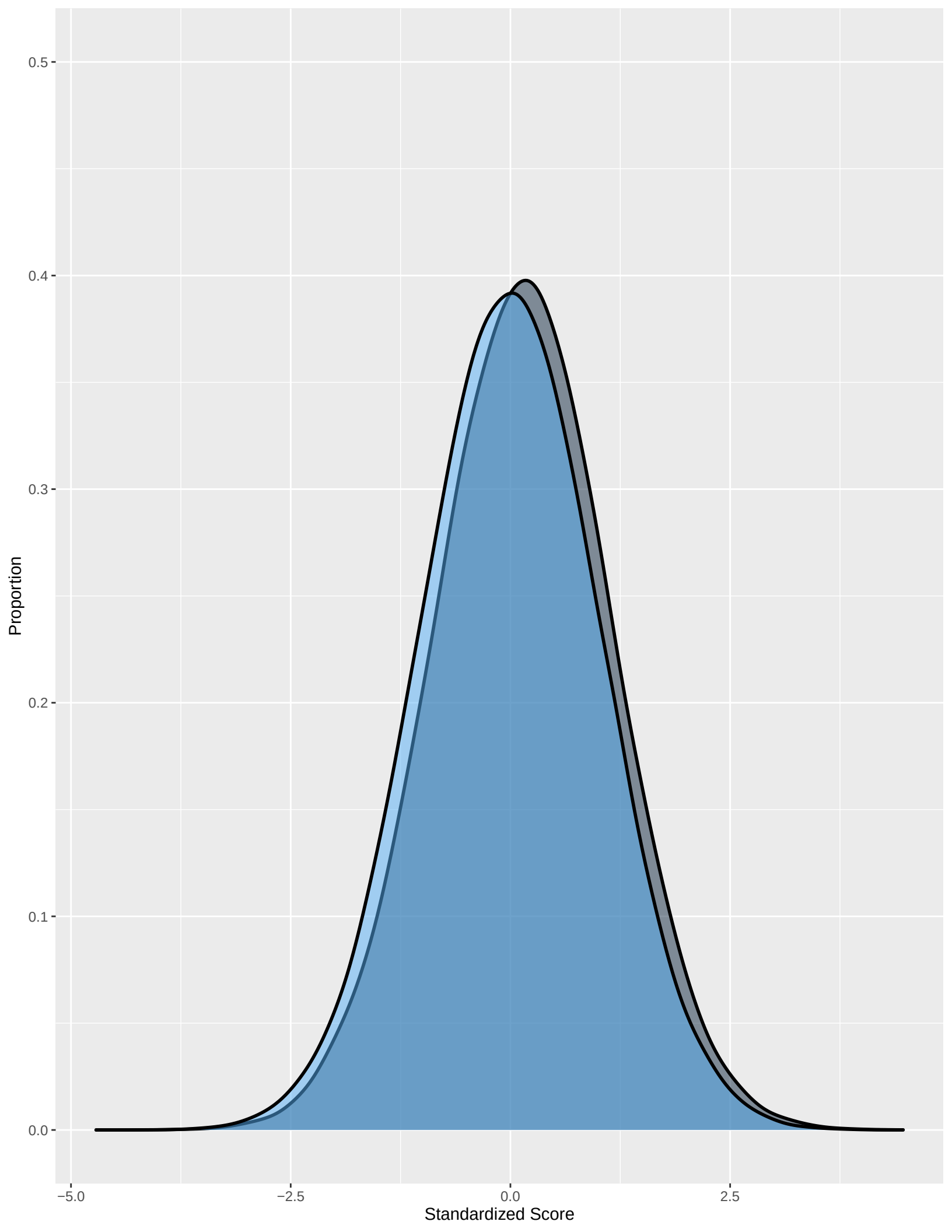
Study	Females/Women Total	Mean	SD	Males/Men Total	Mean	SD	SMD	95%-CI
Savage, 2020	154	3.00	7.00	60	4.00	6.52	-0.15	[-0.44; 0.15]
Shanahan, 2020	378	0.08	0.96	408	0.17	0.86	-0.10	[-0.24; 0.04]
Saraswathi, 2020	139	1.93	7.61	78	2.13	8.03	-0.03	[-0.30; 0.25]
Random effects model	**671**			**546**			**-0.09**	**[-0.21; 0.02]**

Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.84$



Study	Females/Women N	Males/Men N	Difference in Changes	Difference	95% CI
Saraswathi, 2020	139	78		0.04	[-0.11; 0.18]




0.5
0.4
0.3
0.2
0.1
0.0
Proportion
−5.0
−2.5
0.0
2.5
Standardized Score