Comparison of mental health symptoms before and during the covid-19 pandemic: evidence from a systematic review and meta-analysis of 134 cohorts

Ying Sun,1 Yin Wu,1,2 Suqiqiong Fan,1 Tiffany Dal Santo,1,2 Letong Li,1 Xiaowen Jiang,1 Kexin Li,1 Yutong Wang,1 Amina Tasleem,1 Ankur Krishnan,1 Chen He,1 Olivia Bonardi,1 Jill T Boruff,3 Danielle B Rice,4,5 Sarah Markham,6 Brooke Levis,1 Marleine Azar,1 Ian Thombs-Vite,1 Dipika Neupane,1 Branka Agic,7,8 Christine Fahim,9 Michael S Martin,10,11 Sanjeev Sockalingam,7,12 Gustavo Turecki,2,13 Andrea Benedetti,14,15,16 Ying Sun,1 Yin Wu,1,2 Suqiqiong Fan,1 Tiffany Dal Santo,1,2 Letong Li,1 Xiaowen Jiang,1 Kexin Li,1 Yutong Wang,1 Amina Tasleem,1 Ankur Krishnan,1 Chen He,1 Olivia Bonardi,1 Jill T Boruff,3 Danielle B Rice,4,5 Sarah Markham,6 Brooke Levis,1 Marleine Azar,1 Ian Thombs-Vite,1 Dipika Neupane,1 Branka Agic,7,8 Christine Fahim,9 Michael S Martin,10,11 Sanjeev Sockalingam,7,12 Gustavo Turecki,2,13 Andrea Benedetti,14,15,16 Brett D Thombs12,14,15,17,18

ABSTRACT
OBJECTIVE
To synthesise results of mental health outcomes in cohorts before and during the covid-19 pandemic.

DESIGN
Systematic review.

DATA SOURCES
Medline, PsycINFO, CINAHL, Embase, Web of Science, China National Knowledge Infrastructure, Wanfang, medRxiv, and Open Science Framework Preprints.

ELIGIBILITY CRITERIA FOR SELECTING STUDIES
Studies comparing general mental health, anxiety symptoms, or depression symptoms assessed from 1 January 2020 or later with outcomes collected from 1 January 2018 to 31 December 2019 in any population, and comprising ≥90% of the same participants before and during the covid-19 pandemic or using statistical methods to account for missing data. Restricted maximum likelihood random effects meta-analyses (worse covid-19 outcomes representing positive change) were performed. Risk of bias was assessed using an adapted Joanna Briggs Institute Checklist for Prevalence Studies.

RESULTS
As of 11 April 2022, 94 411 unique titles and abstracts including 137 unique studies from 134 cohorts were reviewed. Most of the studies were from high income (n=105, 77%) or upper middle income (n=28, 20%) countries. Among general population studies, no changes were found for general mental health (standardised mean difference (SMD) change 0.11, 95% confidence interval −0.00 to 0.22) or anxiety symptoms (0.05, −0.04 to 0.13), but depression symptoms worsened minimally (0.12, 0.01 to 0.24). Among women or female participants, general mental health (0.22, 0.08 to 0.39), anxiety symptoms (0.20, 0.12 to 0.29), and depression symptoms (0.22, 0.05 to 0.40) worsened by minimal to small amounts. In 27 other analyses across outcome domains among subgroups other than women or female participants, five analyses suggested that symptoms worsened by minimal or small amounts, and two suggested minimal or small improvements. No other subgroup experienced changes across all outcome domains. In three studies with data from March to April 2020 and late 2020, symptoms were unchanged from pre-covid-19 levels at both assessments or increased initially then returned to pre-covid-19 levels. Substantial heterogeneity and risk of bias were present across analyses.

CONCLUSIONS
High risk of bias in many studies and substantial heterogeneity suggest caution in interpreting results. Nonetheless, most symptom change estimates for general mental health, anxiety symptoms, and depression symptoms were close to zero and not statistically significant, and significant changes were of minimal to small magnitudes. Small negative changes occurred for women or female participants in all domains. The authors will update the results of this systematic review as more evidence accrues, with study results posted online (https://www.depressd.ca/covid-19-mental-health).

REVIEW REGISTRATION
PROSPERO CRD42020179703.

Introduction
Concerns about covid-19 related mental health are substantial,1,3 but the sheer volume of low quality evidence has posed a barrier to evidence synthesis and...
decision making.\textsuperscript{6,6} Vast numbers of cross sectional studies have reported proportions of participants with scores above thresholds on easy-to-administer mental health scales as representing the “prevalence” of mental health problems, without comparisons with scores before the covid-19 pandemic.\textsuperscript{5} These scales are not, however, intended to estimate prevalence—thresholds are typically set for screening and to identify far more people than those who have a mental disorder; thus, proportions above thresholds substantially overestimate prevalence.\textsuperscript{7,11} Nonetheless, many study authors and media stories have concluded that the world’s population is experiencing a covid-19 mental health “pandemic” or “tsunami.”\textsuperscript{12}

Many systematic reviews on covid-19 related mental health have synthesised results from cross sectional studies. Two previous systematic reviews compared pre-covid-19 findings with those during the pandemic.\textsuperscript{13,14} One reviewed 65 studies published up to January 2021 and found a small increase in mental health symptoms in early 2020 (standardised mean difference (SMD) 0.11, 95% confidence interval 0.04 to 0.17).\textsuperscript{13} The other searched for studies up to March 2021, included 43 studies, and reported that combined depression and anxiety symptoms worsened early in the pandemic (SMD 0.39, 95% credible interval 0.03 to 0.76).\textsuperscript{14} Both reviews, however, searched a limited number of English language databases, and many relevant studies have been published since they were conducted.

We are conducting an ongoing series of living systematic reviews\textsuperscript{15} on covid-19 related mental health, including a review of studies that compared mental health during covid-19 with pre-pandemic levels in the same cohort.\textsuperscript{4,5} In the present study, we compared general mental health, anxiety symptoms, and depression symptoms in the general population and other groups during covid-19 with outcomes from the same cohorts before covid-19.

Methods

We registered our series of systematic reviews in PROSPERO, and our protocol is available online (https://osf.io/96csg/).\textsuperscript{16} Results are reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.\textsuperscript{17} Supplementary material 1 describes minor amendments to the protocol.

Eligible studies

We included studies on any population that compared eligible outcomes assessed from 1 January 2018 to 31 December 2019 when China first reported covid-19 to the World Health Organization,\textsuperscript{18} with the same outcomes collected from 1 January 2020 or later. Studies had to report data from cohorts comprising at least 90% of the same participants between pre-covid-19 and pandemic periods or to use statistical methods to account for missing data. We did not include repeated cross sectional surveys or studies with fewer than 100 participants.

Eligible outcomes included continuous scores on validated mental health symptom questionnaires, proportion of participants above a threshold on a validated symptom questionnaire, or proportion of participants meeting criteria for a mental disorder using a validated diagnostic interview. In our ongoing living systematic review, we defined outcomes broadly to include, for example, anxiety symptoms, depression symptoms, general mental health, stress, loneliness, anger, grief, and burnout. In the present report, we included only general mental health (eg, general symptoms, mental health related quality of life), anxiety symptoms, and depression symptoms because few studies reported other outcomes. Results for other outcome domains are available online (https://www.depressd.ca/covid-19-mental-health).

Identification and selection of eligible studies

Using a strategy designed by an experienced health sciences librarian, we searched Medline (Ovid), PsycINFO (Ovid), CINAHL (EBSCO), Embase (Ovid), Web of Science Core Collection: Citation Indexes, China National Knowledge Infrastructure, Wanfang, medRxiv, and Open Science Framework Preprints. Because of the need for rapid evidence early in the pandemic, we did not formally peer review the search strategy; however, covid-19 terms were developed in collaboration with other librarians working on the topic and updated as covid-19 specific subject headings became available (see supplementary material 2). We initially searched from 31 December 2019 to 13 April 2020, then daily until 28 December 2020, and thereafter weekly.

Search results were uploaded into DistillerSR (Evidence Partners, Ottawa, Canada), where we identified and removed duplicate references. Two independent reviewers evaluated titles and abstracts in random order. If either reviewer deemed a study potentially eligible, a full text review was completed—also by two independent reviewers. Any discrepancies were resolved through consensus, with a third investigator consulted as necessary. An inclusion and exclusion coding guide was developed and pre-tested, and team members were trained over several sessions (see supplementary material 3).

Data extraction and synthesis

One reviewer extracted data from each included study using a standardised form in DistillerSR, and a second reviewer validated the data using the DistillerSR Quality Control function. Reviewers extracted publication characteristics (eg, first author, publication year, journal); population characteristics, including eligibility criteria, recruitment method, number of participants, and population group (general population, older adults, young adults, children and adolescents, parents, university students, people with pre-existing medical conditions, people with pre-existing mental health conditions, medical staff, and groups defined by sex or gender in the present report, although we extracted any group for which we found data); mental health
Box 1: Interpreting standardised mean differences effect sizes and changes in proportion above a threshold on mental health measures

Changes in symptoms assessed with mental health patient reported outcome measures in covid-19 have been reported as changes in continuous scores and as the proportion of study participants above a threshold. Continuously measured symptom changes are presented in terms of standardised mean difference (SMD), which describe change in terms of within group standard deviations rather than raw change scores, which are measure specific and not easily compared across measures. To illustrate, the figure below shows the amount of change, assuming a normal distribution, for a SMD of 0.25. The hypothetical purple distribution represents pre-covid-19 scores and the orange distribution represents post-covid-19 scores, with a mean symptom increase of SMD 0.25. With a threshold located at 1 standard deviation above the pre-covid-19 mean, the proportion of participants above the threshold would change from 16% to 23%. With a threshold 2 standard deviations above the pre-covid-19 mean, the proportion would change from 2% to 4%.

When studies report an increase or decrease in the proportion of participants above a measure threshold, dichotomous thresholds used for this purpose are sometimes labelled as thresholds for clinically significant symptoms or as reflecting the presence of a condition, such as depression. These designations are not, however, based on evidence that a threshold represents a meaningful divide between impairment and non-impairment and do not reflect the presence of a mental disorder. Most commonly, these designations reflect a point on a measure that balances sensitivity and specificity when used for screening, which does not inform when score levels might become clinically meaningful.

Thresholds on different symptom measures are often located at different places in the symptom distribution. This can lead to divergent estimates of proportions crossing a threshold, depending on the measure used, rather than because of actual differences in symptom changes. As the figure shows, the same change in symptoms in a hypothetical study sample would result in a 7% increase in participants at or above the threshold on one measure (first vertical line, 1 standard deviation above pre-covid-19 distribution mean) but an increase of only 2% on another (second vertical line, 2 standard deviations above pre-covid-19 distribution mean).

outcomes and assessment timing; and adequacy of study methods and reporting. We used World Bank classifications for country income and region. To assess risk of bias and adequacy of study methods and reporting (sampling frame, recruitment methods, sample size, setting and participant descriptions, participation or response rate, outcome assessment methods, standardisation of assessments, statistical analyses, follow-up rate), we adapted the Joanna Briggs Institute Checklist for Prevalence Studies (see supplementary material 4).

For continuous outcomes, we extracted estimates as the SMD effect size with 95% confidence intervals for change from pre-covid-19 to during the pandemic. If such data were not provided, we calculated Hedges g as described elsewhere. Positive SMD estimates represent a worsening of mental health and negative estimates represent an improvement. For proportions, we calculated missing 95% confidence intervals using Newcombe’s method for binomial proportions. For proportion changes, we generated missing 95% confidence intervals using Newcombe’s method for differences between binomial proportions based on paired data, which requires the number of participants above a threshold to be known at both assessment points. If these data were not available, we assumed that 50% of cases above a threshold during pre-covid-19 assessments continued to be above the threshold during the pandemic. We confirmed that results did not differ substantively if we assumed values within a plausible range (30-70%).

Owing to pitfalls in interpreting proportions of participants who crossed a dichotomous threshold, we prioritised continuous data (see box 1). For each population group with continuous outcomes for at least two studies in a domain, we pooled SMDs through restricted maximum likelihood random effects meta-analysis. Heterogeneity was assessed using the I² statistic. For studies with more than one continuous outcome in a domain (eg, two depression symptom measures), we pooled SMDs within the study before fitting the meta-analysis.

Although we have prioritised interpretation of changes in continuous score, in the supplementary material we also report proportions above thresholds, as they can be informative, such as when they are reported for two time points in the same study or as an indicator if some level of change may have occurred.
We have, however, avoided interpretation of the magnitudes of proportions above thresholds. Meta-analyses were performed in R (R version 3.6.3, Rstudio Version 1.2.5042) using the rma.uni function in the metafor package. Forest plots were generated using the forest.rma function in metafor. We characterised changes as minimal (SMD <0.20), small (0.20-0.49), medium (0.50-0.79), or large (0.80-1.99).26

Patient and public involvement
Sarah Markham, an experienced patient advisor and member of The BMJ’s international patient panel, was a research team member from project inception. She provided input on the design of the study, underwent training on study procedures, and was involved in selection of eligible studies. She provided comments on the content of this article.

Results
Search results and study selection
As of 11 April 2022 our searches identified 94,411 unique citations. We excluded 92,457 after review of titles and abstracts and 1523 after full text review, leaving 431 studies with longitudinal data. Of those, 276 studies did not include pre-covid-19 data, 11 only assessed outcomes not included in the present report (e.g., loneliness), one used the same outcome measure but for different time periods pre-covid-19 (worst month in the past year) and during the pandemic (past month), and six duplicated data from another study, leaving 137 unique studies from 134 cohorts (fig 1).

Characteristics of included studies
Supplementary material 5 provides a more detailed version of study characteristics and outcomes. Supplementary table 1 shows the characteristics of the included studies. All cohort studies reported covid-19 outcome data collected in 2020 (nine studies from Asia in January or February, 125 studies from Asia and elsewhere in March or later). Only three cohort studies reported results from multiple time points in 2020, and one also reported results from 2021. Overall, 105 (77%) studies were from high income countries, 28 (20%) from upper middle income countries, one (1%) study from a mix of high income and upper middle income country samples, three (2%) studies from lower middle income countries, and none from low income countries. Fifty two (38%) studies were from Europe and Central Asia, 46 (34%) from East Asia and the Pacific, 28 (20%) from North America, and 11...
Table 1 | Meta-analyses of continuous general mental health, anxiety symptoms, and depression symptoms by population group

<table>
<thead>
<tr>
<th>No of cohorts (No of participants)</th>
<th>Hedges g SMD* (95% CI)</th>
<th>I² (%)</th>
<th>No of cohorts (No of participants)</th>
<th>Hedges g SMD* (95% CI)</th>
<th>I² (%)</th>
<th>No of cohorts (No of participants)</th>
<th>Hedges g SMD* (95% CI)</th>
<th>I² (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>General population</td>
<td>11 (30,185)</td>
<td>0.11 (−0.00 to 0.22)</td>
<td>97</td>
<td>4 (2632)</td>
<td>0.05 (−0.04 to 0.13)</td>
<td>37</td>
<td>4 (3470)</td>
<td>0.12 (0.01 to 0.24)</td>
</tr>
<tr>
<td>Women or female participants</td>
<td>6 (10,329)</td>
<td>0.22 (0.08 to 0.35)</td>
<td>91</td>
<td>51 (3500)</td>
<td>0.20 (0.12 to 0.29)</td>
<td>41</td>
<td>71 (3851)</td>
<td>0.22 (0.05 to 0.40)</td>
</tr>
<tr>
<td>Men or male participants</td>
<td>6 (11,154)</td>
<td>0.11 (−0.12 to 0.39)</td>
<td>98</td>
<td>4 (1271)</td>
<td>0.07 (−0.01 to 0.14)</td>
<td>0</td>
<td>7 (3905)</td>
<td>0.01 (−0.14 to 0.16)</td>
</tr>
<tr>
<td>Older adults</td>
<td>2 (4,221)</td>
<td>0.16 (−0.07 to 0.39)</td>
<td>96</td>
<td>2 (4,602)</td>
<td>0.05 (−0.16 to 0.27)</td>
<td>95</td>
<td>4 (8043)</td>
<td>0.02 (−0.15 to 0.18)</td>
</tr>
<tr>
<td>University students</td>
<td>645 (6957)</td>
<td>0.00 (−0.17 to 0.17)</td>
<td>95</td>
<td>164 (12,642)</td>
<td>−0.07 (−0.21 to 0.06)</td>
<td>96</td>
<td>194 (26,164)</td>
<td>0.14 (0.01 to 0.26)</td>
</tr>
<tr>
<td>Children and adolescents</td>
<td>16 (11,505)</td>
<td>0.19 (−0.05 to 0.42)</td>
<td>99</td>
<td>8 (1,206,4)</td>
<td>0.02 (−0.12 to 0.16)</td>
<td>96</td>
<td>10 (1,167,9)</td>
<td>0.06 (−0.08 to 0.20)</td>
</tr>
<tr>
<td>Parents</td>
<td>3 (9,312)</td>
<td>0.39 (0.21 to 0.56)</td>
<td>57</td>
<td>1 (147)</td>
<td>0.25 (0.02 to 0.49)</td>
<td>96</td>
<td>5 (16,399)</td>
<td>0.15 (−0.05 to 0.35)</td>
</tr>
<tr>
<td>People with pre-existing medical</td>
<td>12 (65,11)</td>
<td>0.10 (−0.01 to 0.20)</td>
<td>86</td>
<td>11 (5775)</td>
<td>0.08 (−0.04 to 0.21)</td>
<td>89</td>
<td>16 (21,594)</td>
<td>0.04 (−0.04 to 0.12)</td>
</tr>
<tr>
<td>health conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>People with pre-existing mental</td>
<td>2 (457)</td>
<td>−0.22 (−0.35 to −0.09)</td>
<td>0</td>
<td>3 (12,362)</td>
<td>0.12 (−0.11 to 0.35)</td>
<td>80</td>
<td>3 (12,352)</td>
<td>−0.05 (−0.08 to −0.03)</td>
</tr>
<tr>
<td>health conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical staff</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>People from sexual or gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>minority groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

See supplementary material for reference details of studies.

©confidence interval, SMD=standardised mean difference.

*Positive effect sizes indicate worse mental health during covid-19 pandemic compared with pre-covid-19.

†Cancer survivors and matched controls from Rentscher et al. counted as one cohort although shown separately in forest plot.

‡Students from Lithuania and Germany from Truskaitiene-Kucareviene et al. counted as one cohort although shown separately in forest plot.

§Sensitivity analysis conducted with results from Savage et al. from October 2020 instead of April 2020; SMD 0.00 (95% CI −0.17 to 0.17), I²=94.8%.

Sensitivity analysis conducted with results from Henry et al. from September to October 2020 instead of April 2020; SMD 0.05 (95% CI −0.04 to 0.14), I²=76%.

**Sensitivity analysis conducted with results from Henry et al. from March 2021 instead of April 2020; SMD 0.04 (95% CI −0.04 to 0.13), I²=76%.

††Sensitivity analysis conducted with results from Henry et al. from September to October 2020 instead of April 2020; SMD 0.03 (95% CI 0.05 to 0.11), I²=91%.

‡‡Sensitivity analysis conducted with results from Henry et al. from March 2021 instead of April 2020; SMD 0.03 (95% CI 0.05 to 0.11), I²=91%.

SMD change for anxiety symptoms in studies of general populations and subgroups. The estimated SMD change for general mental health in studies of general populations and subgroups. The estimated SMD change for anxiety symptoms in studies of general populations and subgroups. The estimated SMD change for depression symptoms in studies of general populations and subgroups.
populations and subgroups. The estimated change in the general population cohorts was not significant and was close to zero (supplementary figure 2a; four cohorts, n=2632; SMD change 0.05, −0.04 to 0.13; I²=37%). Anxiety symptoms worsened significantly by small amounts among women or female participants (supplementary figure 2b; five cohorts, n=3500; SMD change 0.20, 0.12 to 0.29; I²=41%) and parents (one cohort, n=147; SMD change 0.25, 0.02 to 0.49). Estimates were non-statistically significant and close to zero for all other subgroups. I² ranged from 0% to 41% for the general population, women or female participants, and men or male participants but was higher for the other subgroups (80% to 98%).

**Depression symptoms**

The forest plots in supplementary figures 3a-m show the change in SMDs for depression symptoms in studies of general populations and subgroups. The estimated change in general population cohorts increased statistically significantly by a minimal amount (supplementary figure 3a; four cohorts, n=3470; SMD change 0.12, 0.01 to 0.24; I²=81%). The estimated change also increased significantly by minimal to small amounts among women or female participants (supplementary figure 3b; seven cohorts, n=3851; SMD change 0.22, 0.05 to 0.40, I²=89%), older adults (supplementary figure 3d; seven cohorts, n=7419; SMD change 0.22, 0.06 to 0.38, I²=95%), university students (supplementary figure 3f; 19 cohorts, n=26164; SMD change 0.14, 0.01 to 0.26, I²=98%), and people who self-identified as belonging to a sexual or gender minority group (supplementary figure 3k, three cohorts, n=3741; SMD change 0.19, 0.10 to 0.28; I²=67%). The estimated change improved minimally for people with pre-existing mental health conditions (supplementary figure 3j, three cohorts, n=12352; SMD change −0.05, −0.08 to −0.03; I²=0%). The I² statistic was 0% for people with pre-existing mental health conditions and 67% to 98% in other analyses.

**Discussion**

We reviewed more than 94 000 citations and included 137 studies from 134 cohorts that compared mental health during the covid-19 pandemic with assessments pre-covid-19. All studies assessed covid-19 symptoms during at least one time point in 2020, which in most cases was in the first half of the year. Only three studies assessed symptoms in the first months of the pandemic (March to June 2020) and again in late 2020 (September to November), and only one reported results from 2021.

Many analyses showed substantial heterogeneity, which suggests that point estimates should be interpreted cautiously. Consistency did, however, exist across analyses in that most estimates of symptom changes were close to zero and not statistically significant, and changes that were identified were of minimal to small magnitudes. Among general population studies, we did not find changes in general mental health or anxiety symptoms, and the worsening of depression symptoms was minimal (SMD change 0.12).

Among subgroups, women or female participants were the only group that experienced a worsening of symptoms across outcome domains; all by small amounts (SMD changes 0.20 to 0.22). Depression symptoms worsened by minimal to small amounts for older adults, university students, and people who self-identified as belonging to a sexual or gender minority group, but not for other groups. Although based on small numbers of studies and participants, general mental health (three studies, n=932) and anxiety symptoms (one study, 147 participants) were shown to worsen for parents. General mental health and depression symptoms were shown to improve for people with pre-existing mental health conditions, but these findings were based on only two studies (n=457) for general mental health, and improvement was negligible even though statistically significant for depression symptoms (SMD change 0.05). Among the three cohorts in which changes were assessed separately in early 2020 (March to June) and again later in 2020 (September to November) or in 2021, symptoms from later assessments were shown to be either stable or improve from early 2020.

**Comparison with other studies**

Our finding that mental health was either unchanged or worsened by minimal to small amounts in the general population and subgroups is consistent with results from a more limited systematic review of 65 longitudinal studies from early in the pandemic and somewhat smaller than a second systematic review of 43 studies. We know of only one study that has evaluated mental disorders using validated diagnostic methods. That study, from Norway, which was not eligible for our review, evaluated prevalence of current mental disorders in a series of cross sectional random samples collected from 28 January to 11 March 2020 (n=563, 15.4%, 95% confidence interval 12.5% to 18.8%), 12 March to 31 May 2020 (n=691, 9.0%, 7.1% to 11.4%), 1 June to 31 July 2020 (n=530, 14.3%, 11.5% to 17.5%), and 1 August to 18 September 2020 (n=370, 11.9%, 9.0% to 15.6%). The authors concluded that the prevalence of mental health disorders was stable or slightly decreased across the pandemic.

The largest study on suicide during the pandemic included monthly data from official government sources on suicide occurrences from 21 countries from 1 January 2019 or earlier to 31 July 2020 and found no evidence of a statistically significant increase in any country or region; statistically significant decreases did, however, occur in 12 countries or regions. Results were similar in a subsequent update with data from 33 countries across the first nine to 15 months of the pandemic.

We found that women or female participants experienced small negative changes, in aggregate, for general mental health, anxiety symptoms, and depression symptoms during the early part of the pandemic.
towards women increased during the pandemic. 36 Family and elder care. 31-34 Intimate partner violence overrepresented in healthcare jobs and provide most evidence. 12 “tsunami” descriptor or other similar terms used by the media in response to the COVID-19 pandemic on women or female members of the population is of concern. This is an aggregate result that, even though small, suggests that the disproportionate effect of the pandemic on women or female members of the population has influenced mental health. In terms of vulnerabilities, most single parent families tend to be headed by women, and women earn less and are more likely to live in poverty than men. 31-35 Women are also overrepresented in healthcare jobs and provide most family and elder care. 31-34 Intimate partner violence towards women increased during the pandemic. 36 The small overall change in mental health symptoms suggests that many women have been resilient but that among some an important worsening of symptoms occurred. Indeed, although most of our analyses found no changes or minimal to small negative changes in mental health, they do suggest that the pandemic negatively influenced the mental health of some people, which is consistent with, for example, reports of increased visits for mental health and substance misuse. 37 38

Nonetheless, the patterns of findings from our review, along with evidence on mental health disorders and suicide, converge to suggest that the effects of COVID-19 on mental health are more nuanced than the “tsunami” descriptor or other similar terms used by some investigators and in many media articles. 32 Short news cycles that emphasise bad news, 39 40 anecdotes, and an uncritical reliance on cross sectional studies and unvalidated, difficult-to-interpret survey tools that inquire about mental health and wellbeing during the pandemic among convenience samples might at least partially explain this discrepancy. Illustrating the pitfalls of interpreting studies that ask questions about emotional reactions to COVID-19, separate from mental health, a longitudinal study of 2345 young men and women from Switzerland found that depression symptoms and stress significantly decreased during COVID-19 compared with pre-pandemic levels. 41 The study also reported results from a series of unvalidated single items, administered only during the pandemic, that queried about emotional reactions to the pandemic and specifically assigned COVID-19 as the cause (e.g., “due to COVID-19, I experienced . . . “). These items suggested high levels of distress, which became the focus of the study’s conclusions without mention of unchanged mental health indices. Together with the findings from our systematic review, this suggests that many or most people have experienced different aspects of COVID-19 as highly unpleasant or distressing, that most people have been resilient, and that population level mental health has not changed by large amounts, although for some it has changed negatively.

Policy implications
The lack of evidence of a large scale decline in mental health so far in the context of COVID-19 could be because people are resilient and have made the best of a difficult situation. Indeed, although evidence is limited, data suggest, for example, that suicide generally declines during periods of societal conflict. 42-46 War and pandemics have different characteristics, but in both there is a shared threat and common focus on collective action to tackle that threat.

The minimal changes in mental health that we found could also reflect steps that governments have taken to support mental health. The World Health Organization, other pan-national organisations, and governments across the globe have produced strategies for dealing with mental health and have invested in resources to support public mental health. 47 48 Even in countries where mental health had not been a priority previously, 59 60 It is not known to what degree these efforts have been effective, but it is possible that government action has played an important role. However, the negative changes that we detected in some groups, particularly for women and female cohorts, early in the pandemic, underline the need for continued surveillance to determine the degree to which negative mental health changes may be ongoing and require additional resources.

In terms of research, the results of our study underline gaps in mental health surveillance across countries. Since early in the pandemic, the need for high quality surveys with appropriately representative probabilistic sampling methods and pre-pandemic data has been emphasised. 31 However, we found few examples of mental health surveillance frameworks that generated high quality data based on this type of sampling. Investment in more robust mental health surveillance mechanisms that can be used to identify and deal with mental health needs at all times, including in times of crisis, is needed.

Strengths and limitations of this review
Strengths of our systematic review include using rigorous best practice methods; searching nine databases, including two Chinese databases; no restrictions on language; and the ability to update findings as evidence emerged via our living systematic review approach (see https://www.depressed.ca/covid-19-mental-health). Our systematic review also has limitations that suggest some caution in interpreting results. Firstly, we did not peer review our search strategy given the urgency of generating mental health data early in the pandemic. An experienced librarian did, however, develop the search strategy and develop COVID-19 search terms in collaboration with other experienced librarians. We reviewed studies included in other reviews and did not identify any that we
missed; to the contrary, in our review we identified many more studies than other similar reviews with overlapping search periods. Secondly, aside from several population level randomly sampled surveys, most studies included in our systematic review had limitations related to study sampling frames and recruitment methods, response and follow-up rates, and missing follow-up data. Thirdly, we did not include repeated cross sectional studies, which may have excluded some studies with useful evidence. Fourthly, heterogeneity was high in most of the meta-analyses that we conducted. Fifthly, only a handful of studies reported results from late 2020, although the few studies that did suggested that symptoms were stable or reduced from earlier in the pandemic. Sixthly, although we were able to synthesise results from several vulnerable groups, including older adults and people with pre-existing medical conditions, there were few studies for other groups, such as people with low socioeconomic status, and there were no studies on children (ages 0-9 years). Similarly, there was little evidence from low income or lower middle income countries or from some areas of the world, such as sub-Saharan Africa. Seventhly, we did not include studies with fewer than 100 participants; however, in a previous covid-19 mental health systematic review with 65 total studies, studies with fewer than 100 participants comprised only 1% of total participants.

Eighthly, the evidence base is rapidly evolving, and main results could change, although our own living systematic review format will allow updating as this occurs. Finally, we did not assess possible publication bias, although the largely null findings suggest that this was not likely to be an important factor.

Conclusions
We reviewed 137 studies with data from 134 unique cohorts. Across population groups, results suggest that, rather than a mental health crisis, at a population level there has been a high level of resilience during covid-19, and changes in general mental health, anxiety symptoms, and depression symptoms have been minimal to small with no changes detected in most analyses. There were few robust studies with vulnerable groups, however, and it is possible that some population groups experience mental health issues that differ from those of the general population or from other groups. The pandemic and the long term ramifications continue to affect societies across the world, and it will be important to continue to assess mental health. The pandemic has affected the lives of many people, and some are now experiencing mental health difficulties for the first time. Governments should continue to ensure that mental health supports are available and respond to population needs. We will update the results of this systematic review as more evidence accrues, with study results posted online (https://www.depressd.ca/covid-19-mental-health).

AUTHOR AFFILIATIONS
1Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada
2Department of Psychiatry, McGill University, Montreal, Quebec, Canada
3Schulich Library of Physical Sciences, Life Sciences, and Engineering, McGill University, Montreal, Quebec, Canada
4Department of Psychology, St Joseph's Healthcare, Hamilton, Ontario, Canada
5Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada
6Department of Biostatistics and Health Informatics, King’s College London, London, UK
7Centre for Addiction and Mental Health, Toronto, Ontario, Canada
8Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
9Li Ka Shing Knowledge Institute, Unity Health Toronto, Toronto, Ontario, Canada
10School of Epidemiology and Public Health, University of Ottawa, Ontario, Canada
11Correctional Service of Canada, Ottawa, Ontario, Canada
12Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada
13McGill Group for Suicide Studies, Douglas Mental Health University Institute, McGill University, Montreal, Quebec, Canada
14Department of Medicine, McGill University, Montreal, Quebec, Canada
15Department of Medicine, McGill University, Montreal, Quebec, Canada
16Respiratory Epidemiology and Clinical Research Unit, McGill University Health Centre, Montreal, Quebec, Canada
17Department of Psychology, McGill University, Montreal, Quebec, Canada
18Biomedical Ethics Unit, McGill University, Montreal, Quebec, Canada

Contributors: YS, YWu, DBR, AB, and BDT conceived and designed the study. JTB designed the database searches. KL and AK carried out the searches. YS, YWu, SF, TDS, LL, Xi, KL, YWang, AT, AK, CH, OB, DBR, SM, MA, ITV, EK, EW, AY, and BDT contributed to evaluation of study eligibility and data extraction. YS and SF coordinated the study. YS, YWu, BL, AB, and BDT analysed the data. BA, CF, MSM, SS, and GT contributed to interpretation of results as knowledge translation partners. YS 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. The corresponding author and attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This study was funded by the Canadian Institutes of Health Research (CMS-171703, MS1-173070, GA4-177758, WI2-179944) and McGill Interdisciplinary Initiative in Infection and Immunity Emergency covid-19 Research Fund (R2-4). YWu and BL were supported by a Fonds de recherche du Québec – Santé (FRQS) post-doctoral training fellowship. TDS was supported by a Canadian Institutes of Health Research Masters Award. DBR was supported by a Vanier Canada Graduate Scholarship. AB was supported by FRQS senior researcher salary awards, and BDT was supported by a Tier 1 Canada Research Chair, all outside of the present 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.

Competing interests: All authors have completed the ICJME uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: support from the Canadian Institutes of Health Research and McGill Interdisciplinary Initiative in Infection and Immunity Emergency covid-19 Research Fund; 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. BDT and AB declared that they were authors of an included study (see supplementary reference S118).

Ethical approval. Not required.

Data sharing: Study data and statistical codes are available via the Open Science Framework (https://osf.io/96c3g/).
The manuscript’s guarantor (BDT) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned and registered have been explained.

Dissemination to participants and related patient and public communities: Dissemination of findings will be facilitated via posting to the project website (https://www.depressd.ca/covid-19-mental-health), including an infographic with study results; other Canadian and international mental health and covid-19 knowledge exchange portals; a companion article from a patient partner; social media posting; a press release; and direct communication with members of governments.

Provenance and peer review: Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work on means.Introduction to meta-analysis. Wiley & Sons, 2009. doi: 10.1002/9780470333684.ch4.


**Supplementary information:** Supplementary material 1-5, tables 1-5, figures 1-3, and references of included studies