An umbrella review of the benefits and risks associated with youths' interactions with electronic screens

Taren Sanders*[1], Michael Noetel[2], Philip Parker[1], Borja Del Pozo Cruz[3, 14, 15], Stuart Biddle[4, 13], Rimante Ronto[5], Ryan Hulteen[6], Rhiannon Parker[7], George Thomas[8], Katrien De Cocker[9], Jo Salmon[10], Kylie Hesketh[10], Nicole Weeks[1], Hugh Arnott[1], Emma Devine[11], Roberta Vasconcellos[1], Rebecca Pagano[12], Jamie Sherson[12], James Conigrave[1], & Chris Lonsdale[1]

[1] Institute for Positive Psychology and Education, Australian Catholic University, North Sydney, Australia

[2] School of Psychology, University of Queensland, Brisbane, Australia

[3] Department of Sport Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

[4] Centre for Health Research, University of Southern Queensland, Springfield, Australia

[5] Department of Health Sciences, Faculty of Medicine, Health and Human Sciences, Macquarie University, Macquarie Park, Australia

[6] School of Kinesiology, Louisiana State University, Baton Rouge, USA

[7] The Centre for Social Impact, University of New South Wales, Sydney, Australia

[8] The University of Queensland, Health and Wellbeing Centre for Research Innovation, School of Human Movement and Nutrition Sciences, Brisbane, Australia

[9] Department of Movement and Sport Science, Ghent University, Ghent, Belgium

[10] Institute for Physical Activity and Nutrition, Deakin University, Geelong, Australia

[11] The Matilda Centre for Research in Mental Health and Substance Use, University of Sydney, Sydney, Australia

[12] School of Education, Australian Catholic University, North Sydney, Australia

[13] Faculty of Sport & Health Sciences, University of Jyväskylä, Finland

[14] Department of Physical Education, Faculty of Education, University of Cádiz, Cádiz, Spain

[15] Biomedical Research and Innovation Institute of Cádiz (INiBICA) Research Unit, Puerta del Mar University Hospital, University of Cádiz, Cádiz, Spain

Author Note

Correspondence concerning this article should be addressed to Taren Sanders*, 33 Berry St, North Sydney, NSW, Australia. E-mail: Taren.Sanders@acu.edu.au

Abstract

The influence of electronic screens on children and adolescents' health and education is not well understood. In this prospectively registered umbrella review (PROSPERO; CRD42017076051), we harmonised effects from 102 meta-analyses (2,451 primary studies; 1,937,501 participants) on screen time and outcomes. 43 effects from 32 meta-analyses met our criteria for statistical certainty. Meta-analyses of associations between screen use and outcomes showed small-to-moderate effects (range: r = -0.14-0.33). In education, results were mixed; for example, screen use was negatively associated with literacy (r = -0.14, 95% confidence interval [CI] -0.20 to -0.09, p = <0.001, k = 38, N = 18,318), but this effect was positive when parents watched with their children (r = 0.15, 95% CI 0.02 to 0.28, p = 0.028, k = 12, N = 6,083). In health, we found evidence for several small negative associations; for example, social media was associated with depression (r = 0.12, 95% CI 0.05 to 0.19, p = <0.001, k = 12, N = 93,740). Limitations include a limited number of studies for each outcome, medium-to-high risk of bias in 95/102 included meta-analyses and high heterogeneity (17/22 in education and 20/21 in health with $I^2 > 50\%$) We recommend that caregivers carefully weigh the potential harms and benefits of specific types of screen use.

Keywords:

Word count: 5300

An umbrella review of the benefits and risks associated with youths' interactions with electronic screens

Introduction

In the 16th century, hysteria reigned around a new technology that threatened to be "confusing and harmful" to the mind. The cause of such concern? The widespread availability of books brought about by the invention of the printing press.[1] In the early 19th century, concerns about schooling "exhausting the children's brains" followed, with the medical community accepting that excessive study could be a cause of madness.[2] By the 20th century, the invention of the radio was accompanied by assertions that it would distract children from their reading (which by this point was no longer considered confusing and harmful) leading to impaired learning.[3]

Today, the same arguments that were once levelled against reading, schooling, and radio are being made about screen use (e.g., television, mobile phones, and computers).[4] Excessive screen use is the number one concern parents in Western countries have about their children's health and behaviour, ahead of nutrition, bullying, and physical inactivity.[5] Yet, the evidence to support parents' concerns is inadequate. A Lancet editorial[6] suggested that, "Our understanding of the benefits, harms, and risks of our rapidly changing digital landscape is sorely lacking."

While some forms of screen use (e.g., television viewing) may be detrimental to health and wellbeing,[7,8] evidence for other forms of screen exposure (e.g., video games or online communication, such as Zoom™) remains less certain and, in some cases, may even be beneficial.[9,10] Thus, according to a Nature Human Behaviour editorial, research to determine the effect of screen exposure on youth is "a defining question of our age".[11] With concerns over the impact of screen use including education, health, social development, and psychological well-being, an overview that identifies potential benefits and risks is needed.

Citing the negative effects of screens on health (e.g., increased risk of obesity) and health-related behaviours (e.g., sleep), guidelines from the World Health Organisation[12] and numerous government agencies[13,14] and statements by expert groups[15] have recommended that young people's time spent using electronic media devices for entertainment purposes should be limited. For example, the Australian Government guidelines regarding sedentary behaviour recommend that young children (under the age of two) should not spend any time watching screens. They also recommend that children aged 2-5 years should spend no more than one hour engaged in recreational sedentary screen use per day, while children aged 5-12 and adolescents should spend no more than two hours. However, recent evidence suggests that longer exposures may not have adverse effects on children's behaviour or mental health—and might, in fact, benefit their well-being—as long as exposure does not reach extreme levels (e.g., 7 hours per day)[16]. Some research also indicates that content (e.g., video games vs television programs) plays an important role in determining the potential benefit or harm of youths' exposure to screen-based media.[17] Indeed, educational screen use is positively related to educational outcomes.[18] This evidence has led some researchers to argue that a more nuanced approach to screen use guidelines is required.[19]

In 2016, the American Academy of Pediatrics used a narrative review to examine the benefits and risks of children and adolescents' electronic media[20] as a basis for updating their guidelines about screen use.[15] Since then, a large number of systematic reviews and meta-analyses have provided evidence about the potential benefits and risks of screen use. While there have been other overviews of reviews on screen use, these have tended to focus on a single domain (e.g., health[21]), focus on a particular exposure (e.g., social media[22,23]) or provide only a narrative summary of the literature.[24] Focusing on a single domain or exposure makes it difficult to understand what trade-offs are involved in any guidelines around screen use. For example, prohibiting screen use might reduce exposure to advertising but may also thwart learning opportunities from interactive educational tools. Reviews on either of these exposures or outcomes would likely miss being able to quantify these trade-offs. Overviews are one method of evidence synthesis that helps address these trade-offs, by providing 'user-friendly' summaries of a field of research.[25] These overviews provide a reference point for the field and allow for easier comparison of risks and benefits for the same behaviour. By analogy, reading is a sedentary behaviour, and only by comparing the health risks against the educational benefits can researchers and policymakers make clear recommendations about what young people should do.

In order to synthesise the evidence and support further evidence-based guideline development and refinement, we reviewed published meta-analyses examining the effects of screen use on children and youth. This review synthesises evidence on any outcome of electronic media exposure. We deliberately did not pre-specify outcomes, in order to get a list of areas where there is meta-analytical evidence. Adopting this broad approach allowed us to provide a holistic perspective on the influence of screens on children's lives. By synthesising across life domains (e.g., school and home), this review provides evidence to inform guidelines and advice for parents, teachers, pediatricians and other professionals in order to maximise human functioning.

Results

The searches yielded 50,649 results, of which 28,675 were duplicates. After screening titles and abstracts, we assessed 2,557 full-texts for inclusion. Of those, 217 met the inclusion criteria and we extracted the data from all of these meta-analyses. Figure 1 presents the full results of the selection process.

The most frequently reported exposures were physically active video games ($n = 31$), general screen use ($n = 27$), general TV programs and movies ($n = 20$), and screen-based interventions to promote health ($n = 14$). Supplementary File 1 provides a list of all exposures identified. The most frequently reported outcomes were body composition ($n = 30$), general learning ($n = 24$), depression ($n = 13$), and general literacy ($n = 12$). Of the 273 unique exposure/outcome combinations, 241 occurred in only one review, with 23 appearing twice, and 9 appearing three or more times. Full characteristics of the included studies are provided in Supplementary File 2. After removing reviews with duplicate exposure/outcome combinations, our process yielded 252 unique effect/outcome combinations (retaining multiple effects for different age groups or study designs) contributed from 102 reviews. These effects represent the findings of 2,451 primary studies, involving 1,937,501 participants. The characteristics of the included effects are available in Supplementary File 3.

TABLE 1

The quality of the included meta-analyses was mixed (see Table 1). Most assessed heterogeneity (n low risk = 93/102, 91% of meta-analyses), reported the characteristics of the included studies (n low risk = 86/102, 84%), and used a comprehensive and systematic search strategy (n low risk = 71/102, 70%). Most reviews did not clearly report if their eligibility criteria were predefined (n unclear = 71/102, 70%). Many papers also did not complete dual independent screening of abstracts and full text (n high risk = 20/102, 20%) or did not clearly report the method of screening (n unclear = 37/102, 36%). A similar trend was observed for dual independent quality assessment (n high risk = 52/102, 51%; n high risk = 19/102, 19%). Overall, only 7 meta-analyses were graded as low risk of bias on all criteria.

Education Outcomes

There were 88 unique effects associated with education outcomes, including general learning outcomes, literacy, numeracy, and science. We removed 28 effects that did not provide individual study-level data, 19 effects with samples < 1,000, and 19 effects with a significant Egger's test or insufficient studies to conduct the test. Effects not meeting one or more of these standards are presented in Supplementary File 4. The remaining 22 effects met our criteria for statistical credibility and are described in Figure 2. These 22 effects came from 17 meta-analytic reviews analysing data from 337 empirical studies with 262,497 individual participants.

Among the statistically credible effects, general screen use (r = -0.11, 95% confidence interval [CI] -0.24 to 0.01, p = 0.071, k = 18, N = 13,100), television viewing (r = -0.10, 95% CI -0.15 to -0.04, p = <0.001, k = 18, N = 62,135), and video games (r = -0.08, 95% CI -0.12 to -0.04, p = <0.001, k = 10, N = 4,276) were all negatively associated with learning. E-books that included narration (r = 0.11, 95% CI 0.05 to 0.17, p = <0.001, k = 50, N = 2,288), as well as touch screen education interventions (r = 0.21, 95% CI 0.15 to 0.28, p = <0.001, k = 79, N = 5,810), and augmented reality education interventions (r = 0.33, 95% CI 0.25 to 0.42, p = <0.001, k = 15, N = 1,474) were positively associated with learning. General screen use was negatively associated with literacy outcomes (r = -0.14, 95% CI -0.20 to -0.09, p = <0.001, k = 38, N = 18,318). However, if the screen use involved co-viewing (e.g., watching with a parent; r = 0.15, 95% CI 0.02 to 0.28, p = 0.028, k = 12, N = 6,083), or the content of television programs was educational (r = 0.13, 95% CI 0.03 to 0.23, p = 0.012, k = 13, N = 1,955), the association with literacy was positive and significant at the 95% confidence level (weak evidence). Numeracy outcomes were positively associated with screen-based mathematics interventions (r = 0.27, 95% CI 0.21 to 0.33, p =

<0.001, $k = 85$, $N = 36{,}793$) and video games that contained numeracy content ($r = 0.32$, 95% CI 0.21 to 0.43, $p = <0.001$, $k = 25$, $N = 2{,}008$).

As shown in Figure 2, most of the credible results (13 of 22 effects) showed statistically significant associations, with 99.9% confidence intervals not encompassing zero (strong evidence). The remaining six associations were significant at the 95% confidence level (weak evidence). All credible effects related to education outcomes were small-to-moderate. Screen-based interventions designed to influence an outcome (e.g., a computer based program designed to enhance learning;[26] $r = 0.21$, 95% CI 0.15 to 0.28, $p = <0.001$, $k = 79$, $N = 5{,}810$) tended to have larger effect sizes than exposures that were not specifically intended to influence any of the measured outcomes (e.g., the association between television viewing and learning;[27] $r = -0.10$, 95% CI -0.15 to -0.04, $p = <0.001$, $k = 18$, $N = 62{,}135$). The largest effect size observed was for augmented reality-based education interventions on general learning ($r = 0.33$, 95% CI 0.25 to 0.42, $p = <0.001$, $k = 15$, $N = 1{,}474$). Most effects showed high levels of heterogeneity (17 of 22 with $I^2 > 50\%$).

Health-related Outcomes

We identified 163 unique outcome-exposure combinations associated with health or health-related behaviour outcomes. We removed 39 effects that did not provide individual study-level data, 50 effects with samples $< 1{,}000$, and 53 effects with a significant Egger's test or insufficient studies to conduct the test. No remaining studies had statistically significant tests for excess significance. Effects not meeting one or more of these standards are presented in Supplementary File 5. The remaining 21 meta-analytic associations met our criteria for credible evidence and are described below (see also Figure 3). These 21 effects came from 15 meta-analytic reviews analysing data from 344 empirical studies with 859,562 individual participants.

Digital advertising of unhealthy foods—both traditional advertising ($r = 0.23$, 95% CI

0.10 to 0.37, p = <0.001, k = 13, N = 1,756) and video games developed by a brand for promotion (r = 0.18, 95% CI 0.10 to 0.25, p = <0.001, k = 15, N = 3,842)—were associated with higher unhealthy food intake. Social media use and sexual content were positively associated with risky behaviors (e.g., social media and risky sexual behaviour; r = 0.21, 95% CI 0.14 to 0.28, p = <0.001, k = 14, N = 23,096). Television viewing was negatively correlated with sleep duration, but with stronger evidence only observed for adolescents (r = -0.06, 95% CI -0.10 to -0.01, p = 0.018, k = 10, N = 9,798). Both television and video games were associated with body composition (e.g., television r = 0.06, 95% CI 0.03 to 0.10, p = <0.001, k = 12, N = 3,196). Screen-based interventions which target health behaviours appeared mostly effective.

Across the health outcomes, most (14 of 21) effects were statistically significant at the 99.9% confidence interval level, with the remaining four significant at 95% confidence. However, most of the credible effects exhibited high levels of heterogeneity, with all but two having $I^2 > 75\%$. Additionally, most effects were small, with the association between internet use and depression the largest at r = 0.25 (95% CI 0.22 to 0.27, p = <0.001, k = 118, N = 527,696). Most of the effect sizes (17/21) had an absolute value of $r < 0.2$.

Discussion

The primary goal of this review was to provide a holistic perspective on the influence of screens on children's lives across a broad range of outcomes. We found that when meta-analyses examined general screen use, and did not specify the content, context or device, there was strong evidence showing potentially harmful associations with general learning, literacy, body composition, and depression. However, when meta-analyses included a more nuanced examination of exposures, a more complex picture appeared.

As an example, consider children watching television programs—an often cited form of screen use harm. We found evidence for a small association with poorer academic performance and literacy skills for general television watching[27]. However, we also found evidence that if the content of the program was educational, or the child was watching the program with a parent (i.e., co-viewing), this exposure was instead associated with better literacy.[28] Thus, parents may play an important role in selecting content that is likely to benefit their children or, perhaps, interact with their children in ways that may foster literacy (e.g., asking their children questions about the program). Similar nuanced findings were observed for video games. The credible evidence we identified showed that video game playing was associated with poorer body composition and learning.[27,29] However, when the video game were designed specifically to teach numeracy, playing these games showed learning benefits.[30] One might expect that video games designed to be physically active could confer health benefits, but none of the meta-analyses examining this hypothesis met our thresholds for statistical credibility (see Supplementary Files 4 & 5) therefore this hypothesis could not be addressed.

Social media was one type of exposure that showed consistent associations with poor health, with no indication of potential benefit. Social media showed strong evidence of harmful associations with risk taking in general, as well as unsafe sex and substance abuse.[31] These results align with meta-analytic evidence from adults indicating that social media use is also associated with increased risk of depression.[32,33] Recent evidence from social media companies themselves suggest there may also be negative effects of social media on the mental health of young people, especially teenage girls.[34]

One category of exposure appeared to be consistently associated with benefits: screen-based interventions designed to promote learning or health behaviours. This finding indicates that interventions can be effectively delivered using electronic media platforms, but does not necessarily indicate that screens are more effective than other methods (e.g., face-to-face, printed material). Rather, it reinforces that the content of the screen use may be the most important aspect. The way that a young person interacts with digital screens may also be important. We found evidence that touch screens had strong evidence for benefits on learning,[26] as did augmented reality.[35]

Largely owing to a small number of studies or missing individual study data, there were few age-based conclusions that could be drawn from reviews which met our criteria for statistical certainty. Given the differences in development across childhood and adolescence and the different ways children of various ages use screens, further examination of age-based differences is needed. However, in the absence of this work, our study has shown how children are affected by screens in general.

Among studies that met our criteria for statistical certainty heterogeneity was high, with almost all effects having $I^2 > 50\%$. Much of this heterogeneity is likely explained by differences in measures across pooled studies, or in some cases, the generic nature of some of the exposures. For example, “TV programs and movies” covers a substantial range of content, which may explain the heterogeneous association with education outcomes.

Our results have several implications for policy and practice. Broadly, our findings align with the recommendations of others who suggest that current guidelines may be too simplistic, mischaracterise the strength of the evidence, or do not acknowledge the important nuances of the issue.[36–38] Our findings suggest that screen use is a complex issue, with associations based not just on duration and device type, but also on the content and the environment in which the exposure occurs. Many current guidelines simplify this complex relationship as something that should be minimised.[12,13] We suggest that future guidelines need to embrace the complexity of the issue, to give parents and clinicians specific information to weigh the pros and cons of interactions with screens.

Given our results, we support the continuing trend of guidelines moving away from recommendations to reduce ‘screen use’, and instead focusing on the type of screen use. For example, we suggest that guidelines should discourage high levels of social media and internet use. Guidelines may also consider adapting recommendations that promote the use of educational apps and video games, although these recommendations need to be balanced against the (very small) risks to adiposity.[39]

Our results also have implications for future research. Screen use research is extensive, varied, and rapidly growing. Reviews tended to be general (e.g., all screen use) and even when more targeted (e.g., social media) nuances related to specific content (e.g., Instagram vs Facebook) have not been meta-analysed or have not produced credible evidence. Fewer than 20% of the effects identified met our criteria for statistical credibility. Most studies which did not meet our criteria failed to provide study-level data (or did not provide sufficient data, such as including effect estimates but not sample sizes). Newer reviews were more likely to provide this information than older reviews, but it highlights the importance of data and code sharing as recommended in the PRISMA guidelines.[40] When study level data was available, many effects were removed because the pooled sample size was small, or because there were fewer than ten studies on which to perform an Egger's test. It seems that much of the current screen use research is small in scale, and there is a need for larger, high-quality studies.

Our results highlight the need for the field to more carefully consider if the term 'screen use' remains appropriate for providing advice to parents. Instead, our results suggest that more nuanced and detailed descriptions of the behaviours to be modified may be required. Rather than suggesting parents limit 'screen use', for example, it may be better to suggest that parents promote interactive educational experiences but limit exposure to advertising.

Screen use research has a well-established measurement problem, which impacts the individual studies of this umbrella review. The vast majority of screen use research relies on self-reported data, which not only lacks the nuance required for understanding the effects of screen use, but may also be inaccurate. In one systematic review on screen use and sleep,[7] 66 of the 67 included studies used self-reported data for *both* the exposure and outcome variable.

It has been established that self-reported screen use data has questionable validity. In a meta-analysis of 47 studies comparing self-reported media use with logged measures, Parry et al[41] found that the measures were only moderately correlated ($r = 0.38$), with self-reported problematic usage fairing worse ($r = 0.25$). Indeed, of 622 studies which measured the screen use of 0—6 year-olds, only 69 provided any sort of psychometric properties for their measure, with only 19 studies reporting validity.[42] While some researchers have started using newer methods of capturing screen behaviours—such as wearable cameras[43] or device-based loggers[44]—these are still not widely adopted. It may be that the field of screen use research cannot be sufficiently advanced until accurate, validated, and nuanced measures are more widely available and adopted.

There were a number of strengths and limitations to our work. Our primary goal for this umbrella review was to provide a high-level synthesis of screen use research, by examining a range of exposures and the associations with a broad scope of outcomes. Our results represent the findings from 2,451 primary studies comprised of 1,937,501 participants. To ensure findings could be compared on a common metric, we extracted and reanalysed individual study data where possible.

Our high-level approach limits the feasibility of examining fine-grained details of the individual studies. For example, we did not examine moderators beyond age, nor did we rate the risk of bias for the individual studies. Thus, our assessment of evidence quality was restricted to statistical credibility, rather than a more complete assessment of quality (e.g., GRADE[45]). As such, we made decisions regarding the credibility of evidence, where others may have used different thresholds or metrics. In addition, when faced with duplicate outcome/exposure combinations we chose to keep the one with the largest pooled sample size, assuming that this would capture the most comprehensive and most recent review. Inspection of the excluded effect sizes suggests that this decision was not that impactful: our results would have been almost exactly the same has we used the number of included studies

(k) or the most recent review by publication year. However, we provide the complete results in Supplementary Files 4 & 5, along with the dataset (Supplementary File 6) for others to consider alternative criteria.

Our high-level approach also means that we could not engage with the specific mechanisms behind each association, and as such, we cannot make claims on the directions of causality. These likely depend on the specific exposure and outcome. It is tempting to draw inferences that the associations are due to screen use causing these outcomes, but we cannot rule out reverse causality, a third variable, or some combination of influences. Many of the individual reviews go into more detail about the strength of the evidence for causal associations, but those judgements were difficult to synthesise across more than 200 reviews. Readers who wish to more deeply understand one specific relationship are directed to the cited review for that effect, where the authors could engage more deeply with the mechanisms.

We converted all effect sizes to a common metric (Pearson's r) to allow for comparisons of magnitude, but acknowledge that this assumes a linear relationship between the variables. Some previous research suggests that associations are typically linear.[18] However, others have identified instances where non-linear relationships exist, especially for very high levels of screen use.[17,46,47] Additionally, our conversion may not always adequately account for differences in study design or measures of exposures and outcomes. Care is needed, therefore, when interpreting the effect sizes. In addition, reviews provide only historical evidence which may not keep up with the changing ways children can engage with screens. While our synthesis of the existing evidence provides information about how screens might have influenced children in the past, it is difficult to know if these findings will translate to new forms of technology in the future.

Screen use is a topic of significant interest, as shown by the wide variety of academic domains involved, parents' concerns, and the growing pervasiveness into society. Our findings showed that screen use is associated with both positive (e.g., educational video games were associated with improved literacy) and negative (e.g., general screen use was associated with poorer body composition) outcomes. Based on our findings, we recommend that parents, teachers, and other caregivers need to carefully weigh the pros and cons of each specific activity for potential harms and benefits. However, our findings also lead us to suggest that in order to aid caregivers to make this judgement, researchers need to conduct more careful and nuanced measurement and analysis of screen use, with less emphasis on measures that aggregate screen use and instead focus on the content, context, and environment in which the exposure occurs.

Methods

We prospectively registered our methods on the International Prospective Register of Systematic Reviews (PROSPERO; CRD42017076051) in October 2017. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[40]

Eligibility criteria. Population: To be eligible for inclusion, meta-analyses needed to include meta-analytic effect sizes for children or adolescents (age 0-18 years). We included meta-analyses containing studies that combined data from adults and youth if meta-analytic effect size estimates specific to participants aged 18 years or less could be extracted (i.e., the highest mean age for any individual study included in the meta-analysis was $<$ 18 years). A meta-analysis was still included if the age range exceed 18 years, provided that the mean age was less than 18. We excluded meta-analyses that only contained evidence gathered from adults (age $>$18 years).

Exposure: We included meta-analyses examining all types of electronic screens including (but not necessarily limited to) television, gaming consoles, computers, tablets, and mobile phones. We also included analyses of all types of content on these devices, including (but not necessarily limited to) recreational content (e.g., television programs, movies, games), homework, and communication (e.g., video chat). In this review we focused on electronic media exposure that would be considered typical for children and youth. That is, exposure that may occur in the home setting, or during schooling. Consistent with this approach, we excluded technology-based treatments for clinical conditions. However, we included studies examining the effect of screen exposure on non-clinical outcomes (e.g., learning) for children and youth with a clinical condition. For example, a meta-analysis of the effect of television watching on learning among adolescents diagnosed with depression would be included. However, a meta-analysis of interventions designed to *treat* clinical depression delivered by a mobile phone app would be excluded.

Outcomes: We included all reported outcomes on benefits and risks.

Publications: We included meta-analyses (or meta-regressions) of quantitative evidence. To be included, meta-analyses needed to analyse data from studies identified in a systematic review. For our purposes, a systematic review was one in which the authors attempted to acquire all the research evidence that pertained to their research question(s). We excluded meta-analyses that did not attempt to summarise all the available evidence (e.g., a meta-analysis of all studies from one laboratory). We included meta-analyses regardless of the study designs included in the review (e.g., laboratory-based experimental studies, randomised controlled trials, non-randomised controlled trials, longitudinal, cross-sectional, case studies), as long as the studies in the review collected quantitative evidence. We excluded systematic reviews of qualitative evidence. We did not formulate inclusion/exclusion criteria related to the risk of bias of the review. We did, however, employ a risk of bias tool to help interpret the results. We included full-text, peer-reviewed meta-analyses published or 'in-press' in English. We excluded conference abstracts and meta-analyses that were unpublished.

Information sources. We searched records contained in the following databases: Pubmed, MEDLINE, CINAHL, PsycINFO, SPORTDiscus, Education Source, Embase, Cochrane Library, Scopus, Web of Science, ProQuest Social Science Premium Collection, and ERIC. We conducted an initial search on August 17, 2018 and refreshed the search on September 27, 2022. We searched reference lists of included papers in order to identify additional eligible meta-analyses. We also searched PROSPERO to identify relevant protocols and contacted authors to determine if these reviews have been completed and published.

Search strategy. The search strategy associated with each of the 12 databases can be found in Supplementary File 7. We hand searched reference lists from any relevant umbrella reviews to identify systematic meta-analyses that our search may have missed.

Selection process. Using Covidence software (Veritas Health Innovation, Melbourne, Australia), two researchers independently screened all titles and abstracts. Two researchers then independently reviewed full-text articles. We resolved disagreements at each stage of the process by consensus, with a third researcher employed, when needed.

Data items. From each included meta-analysis, two researchers independently extracted data into a custom-designed database. We extracted the following items: First author, year of publication, study design restrictions (e.g., cross-sectional, observational, experimental), region restrictions (e.g., specific countries), earliest and latest study publication dates, sample age (mean), lowest and highest mean age reported, outcomes reported, and exposures reported.

Study risk of bias assessment. For each meta-analysis, two researchers independently completed the National Health, Lung and Blood Institute's Quality Assessment of Systematic Reviews and Meta-Analyses tool[48] (see Table 1). We resolved disagreements by consensus, with a third researcher employed when needed. We did not assess risk of bias in the individual studies that were included in each meta-analysis.

Effect measures. Two researchers independently extracted all quantitative meta-analytic effect sizes, including moderation results. We excluded effect sizes which were reported as relative risk ratios or odds ratios, as meta-analyses did not contain sufficient information to meaningfully convert to a correlation. We also excluded effect size estimates when the authors did not provide a sample size. Where possible, we also extracted effect sizes from the primary studies included in each meta-analysis.

To facilitate comparisons, we converted effect sizes to Pearson's r using established formulae.[49,50] Effect sizes on the original metric are provided in Supplementary File 6. Throughout the results section we interpret the size of the effects using Funder and Ozer's guidelines:[51] very small ($0.05 < r <= 0.1$), small ($0.1 < r <= 0.2$), medium ($0.2 < r <= 0.2$), large ($0.3 < r <= 0.4$), and very large ($r >= 0.4$). These are similar to other interpretations based on empirical data.[52]

Synthesis methods. After extracting data, we examined the combinations of exposure and outcomes and removed any effects that appeared multiple times (i.e., in multiple meta-analyses, or with multiple sub-groups in the same meta-analysis), keeping the effect with the largest total sample size. In instances where effect sizes from the same combination of exposure and outcome were drawn from different age-groups (e.g., children vs adolescents), or were drawn using different study designs (e.g., cross-sectional vs longitudinal) we retained both estimates in our dataset.

We descriptively present the remaining meta-analytic effect sizes. To remove the differences in approach to meta-analyses across the reviews, we reran the effect size estimate using a random effects meta-analysis via the metafor package[53] in R[54] (version 4.3.0) when the meta-analysis's authors provided primary study data associated with these effects. When required, we imputed missing sample sizes using mean imputation from the other studies within that review. From our reanalysis we also extracted I^2 values. To test for publication bias, we conducted Egger's test[55] when the number of studies within the review was ten or more,[56] and conducted a test of excess significance.[57] We contacted authors who did not provide primary study data in their published article. Where authors did not provide data in a format that could be re-analysed, we used the published results of their original meta-analysis.

Evidence assessment criteria. Statistical Credibility: We employed a statistical classification approach to grade the credibility of the effect sizes in the literature. To be considered 'credible' an effect needed to be derived from a combined sample of >1,000 participants[58] and have non-significant tests of publication bias (i.e., Egger's test and excess significance test). We performed these analyses, and therefore the review needed to provide usable study-level data in order to be included.

Consistency of Effect within the Population: We also examined the consistency of the effect size using the I^2 measure. We considered $I^2 < 50\%$ to indicate effects that were relatively consistent across the population of interest. I^2 values of $> 50\%$ were taken to indicate an effect was potentially heterogeneous within the population.

Direction of Effect: Finally, we examined the extent to which significance testing suggested screen exposure was associated with benefit, harm, or no effect on outcomes. We used thresholds of $P < .05$ for weak evidence (i.e., 95% confidence intervals did not cross zero) and $P < 10^{-3}$ (i.e., 99.9% confidence intervals did not cross zero) for strong evidence. An effect with statistical credibility but with $P > .05$ (i.e., 95% confidence intervals included zero) was taken to indicate no association of interest.

Deviations from protocol. As described above, we have summarised the meta-analytic findings from all included systematic reviews. In our protocol, we originally planned to also conduct a narrative synthesis of all systematic reviews, even those without meta-analyses. However, we determined that combining results from the meta-analyses alone allow readers to compare relative strength of associations more easily. Readers interested in the relevant systematic reviews (i.e., without meta-analysis) can consult the list of references in Supplementary File 8.

We altered our evidence assessment plan when we identified that, as written, it could not classify precise evidence of null effects (i.e., from large reviews with low heterogeneity and low risk of publication bias) as 'credible' because a highly-significant P-value was a criteria. This would have significantly harmed knowledge gained from our review as it would have restricted our ability to show where the empirical evidence strongly indicated that there was no association between screen use and a given outcome.

Data availability statement

All data for this review are available from the authors' GitHub repository (https://github.com/motivation-and-Behaviour/screen_umbrella) or from the Open Science Foundation (https://osf.io/3ubqp/).

Code availability statement

All code used in these analyses are available on the authors' GitHub repository (https://github.com/motivation-and-Behaviour/screen_umbrella).

Acknowledgements

The authors received no specific funding for this work.

Author contributions

TS, MN, PP, and CL conceptualised the review and drafted the manuscript. TS, MN, and PP conducted the analyses. All authors contributed to data extraction, interpretation, and editing of the manuscript.

Competing interests

The authors declare no conflicts of interest.

Tables

Table 1: Review characteristics and quality assessment for meta-analyses providing unique effects

Figure legends

Figure 1: PRISMA flow diagram.

Figure 2: Education outcomes. Forest plot for 22 unique effect sizes related to educational outcomes which met the criteria for statistical certainty. Findings are presented as correlations (two-sided) with both 95% and 99.9% confidence intervals.

Figure 3: Health and health-related behaviour outcomes. Forest plot for 21 unique effect sizes related to health and health-related behaviour outcomes which met the criteria for statistical certainty. Findings are presented as correlations (two-sided) with both 95% and 99.9% confidence intervals.

References

1. Blair, A. Reading Strategies for Coping With Information Overload ca.1550-1700. *Journal of the History of Ideas* **64**, 11–28 (2003).

2. Bell, A. N. *The sanitarian.* vol. 11 (AN Bell, 1883).

3. Dill, K. E. *The Oxford handbook of media psychology.* (Oxford University Press, 2013).

4. Wartella, E. A. & Jennings, N. Children and computers: New technology. Old concerns. *The future of children* 31–43 (2000).

5. Rhodes, A. *Top ten child health problems: What the public thinks.* (2015).

6. The Lancet. Social media, screen time, and young people's mental health. *The Lancet* **393**, 611 (2019).

7. Hale, L. & Guan, S. Screen time and sleep among school-aged children and adolescents: A systematic literature review. *Sleep Medicine Reviews* **21**, 50–58 (2015).

8. Sweetser, P., Johnson, D., Ozdowska, A. & Wyeth, P. Active versus passive screen time for young children. *Australasian Journal of Early Childhood* **37**, 94–98 (2012).

9. Li, X. & Atkins, M. S. Early childhood computer experience and cognitive and motor development. *Pediatrics* **113**, 1715–1722 (2004).

10. Warburton, W. & Highfield, K. Children and technology in a smart device world. in *Children, Families and Communities* 195–221 (Oxford University Press, 2017).

11. Nature Human Behaviour. Screen time: How much is too much? *Nature* **565**, 265–266 (2019).

12. World Health Organization. *Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age.* 33 p. (World Health Organization, 2019).

13. Australian Government. *Physical activity and exercise guidelines for all Australians.* (2021).

14. Canadian Society for Exercise Physiology. *Canadian 24-Hour Movement Guidelines for Children and Youth: An Integration of Physical Activity, Sedentary Behaviour, and Sleep.* (2016).

15. Council On Communication and Media. Media Use in School-Aged Children and Adolescents. *Pediatrics* **138**, e20162592 (2016).

16. Ferguson, C. J. Everything in Moderation: Moderate Use of Screens Unassociated with Child Behavior Problems. *Psychiatric Quarterly* **88**, 797–805 (2017).

17. Przybylski, A. K. & Weinstein, N. A Large-Scale Test of the Goldilocks Hypothesis: Quantifying the Relations Between Digital-Screen Use and the Mental Well-Being of Adolescents. *Psychological Science* **28**, 204–215 (2017).

18. Sanders, T., Parker, P. D., del Pozo-Cruz, B., Noetel, M. & Lonsdale, C. Type of screen time moderates effects on outcomes in 4013 children: Evidence from the Longitudinal Study of Australian Children. *International Journal of Behavioral Nutrition and Physical Activity* **16**, 117 (2019).

19. Kaye, L. K., Orben, A., Ellis, D. A., Hunter, S. C. & Houghton, S. The Conceptual and Methodological Mayhem of 'Screen Time'. *International Journal of Environmental Research and Public Health* **17**, 3661 (2020).

20. Chassiakos, Y. L. R. *et al.* Children and Adolescents and Digital Media. *Pediatrics* **138**, e20162593 (2016).

21. Stiglic, N. & Viner, R. M. Effects of screentime on the health and well-being of children and adolescents: A systematic review of reviews. *BMJ Open* **9**, e023191 (2019).

22. Valkenburg, P. M., Meier, A. & Beyens, I. Social media use and its impact on adolescent mental health: An umbrella review of the evidence. *Current Opinion in Psychology* **44**, 58–68 (2022).

23. Arias-de la Torre, J. *et al.* Relationship Between Depression and the Use of Mobile Technologies and Social Media Among Adolescents: Umbrella Review. *Journal of Medical Internet Research* **22**, e16388 (2020).

24. Orben, A. Teenagers, screens and social media: A narrative review of reviews and key studies. *Social Psychiatry and Psychiatric Epidemiology* **55**, 407–414 (2020).

25. Pollock, M., Fernandes, R., Becker, L., Pieper, D. & Hartling, L. Chapter V: Overviews of Reviews. in *Cochrane Handbook for Systematic Reviews of Interventions* (eds. Higgins, J. P. et al.) (Cochrane, 2022).

26. Xie, H. *et al.* Can Touchscreen Devices be Used to Facilitate Young Children's Learning? A Meta-Analysis of Touchscreen Learning Effect. *Frontiers in Psychology* **9**, 2580 (2018).

27. Adelantado-Renau, M. *et al.* Association Between Screen Media Use and Academic Performance Among Children and Adolescents: A Systematic Review and Meta-analysis. *JAMA Pediatrics* **173**, 1058 (2019).

28. Madigan, S., McArthur, B. A., Anhorn, C., Eirich, R. & Christakis, D. A. Associations Between Screen Use and Child Language Skills: A Systematic Review and Meta-analysis. *JAMA Pediatrics* **174**, 665 (2020).

29. Poorolajal, J., Sahraei, F., Mohamdadi, Y., Doosti-Irani, A. & Moradi, L.

Behavioral factors influencing childhood obesity: A systematic review and meta-analysis. *Obesity Research & Clinical Practice* **14**, 109–118 (2020).

30. Byun, J. & Joung, E. Digital game-based learning for K-12 mathematics education: A meta-analysis. *School Science and Mathematics* **118**, 113–126 (2018).

31. Vannucci, A., Simpson, E. G., Gagnon, S. & Ohannessian, C. M. Social media use and risky behaviors in adolescents: A meta-analysis. *Journal of Adolescence* **79**, 258–274 (2020).

32. Yoon, S., Kleinman, M., Mertz, J. & Brannick, M. Is social network site usage related to depression? A meta-analysis of FacebookDepression relations. *Journal of Affective Disorders* **248**, 65–72 (2019).

33. Vahedi, Z. & Zannella, L. The association between self-reported depressive symptoms and the use of social networking sites (SNS): A meta-analysis. *Current Psychology* **40**, 2174–2189 (2021).

34. Seetharaman, G. W., Jeff Horwitz and Deepa. Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show. *Wall Street Journal* (2021).

35. Tekedere, H. & Göke, H. Examining the Effectiveness of Augmented Reality Applications in Education: A Meta-Analysis. *International Journal of Environmental and Science Education* **11**, 9469–9481 (2016).

36. Elson, M. *et al.* Do policy statements on media effects faithfully represent the science? *Advances in Methods and Practices in Psychological Science* **2**, 12–25 (2019).

37. Ashton, J. J. & Beattie, R. M. Screen time in children and adolescents: Is there evidence to guide parents and policy? *The Lancet Child & Adolescent Health* **3**, 292–294 (2019).

38. Royal College of Paediatrics and Child Health. *The health impacts of screen time: A guide for clinicians and parents.* (2019).

39. Marker, C., Gnambs, T. & Appel, M. Exploring the myth of the chubby gamer: A meta-analysis on sedentary video gaming and body mass. *Social Science & Medicine* **301**, 112325 (2022).

40. Page, M. J. *et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews.* (2020) doi:10.31222/osf.io/v7gm2.

41. Parry, D. A. *et al.* A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour* **5**, 1535–1547 (2021).

42. Byrne, R., Terranova, C. O. & Trost, S. G. Measurement of screen time among young children aged 0 years: A systematic review. *Obesity Reviews* **22**, (2021).

43. Smith, C., Galland, B. C., de Bruin, W. E. & Taylor, R. W. Feasibility of automated cameras to measure screen use in adolescents. *American journal of preventive medicine* **57**, 417–424 (2019).

44. Ryding, F. C. & Kuss, D. J. Passive objective measures in the assessment of problematic smartphone use: A systematic review. *Addictive Behaviors Reports* **11**, 100257 (2020).

45. Guyatt, G. *et al.* GRADE guidelines: 1. IntroductionGRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* **64**, 383–394 (2011).

46. Twenge, J. M. More Time on Technology, Less Happiness? Associations Between Digital-Media Use and Psychological Well-Being. *Current Directions in Psychological Science* **28**, 372–379 (2019).

47. Kelly, Y., Zilanawala, A., Booker, C. & Sacker, A. Social Media Use and

Adolescent Mental Health: Findings From the UK Millennium Cohort Study. *EClinicalMedicine* **6**, 59–68 (2018).

48. National Health, Lung, and Blood Institute. *Quality Assessment of Systematic Reviews and Meta-Analyses.* (2014).

49. Bowman, N. A. Effect Sizes and Statistical Methods for Meta-Analysis in Higher Education. *Research in Higher Education* **53**, 375–382 (2012).

50. Jacobs, P. & Viechtbauer, W. Estimation of the biserial correlation and its sampling variance for use in meta-analysis: Biserial Correlation. *Research Synthesis Methods* **8**, 161–180 (2017).

51. Funder, D. C. & Ozer, D. J. Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science* **2**, 156–168 (2019).

52. Gignac, G. E. & Szodorai, E. T. Effect size guidelines for individual differences researchers. *Personality and Individual Differences* **102**, 74–78 (2016).

53. Viechtbauer, W. *Metafor: Meta-analysis package for r.* (2023).

54. R Core Team. *R: A language and environment for statistical computing.* (R Foundation for Statistical Computing, 2023).

55. Egger, M., Smith, G. D., Schneider, M. & Minder, C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**, 629–634 (1997).

56. Page, M. J., Higgins, J. P. & Sterne, J. A. Chapter 13: Assessing risk of bias due to missing results in a synthesis. in *Cochrane Handbook for Systematic Reviews of Interventions* (eds. Higgins, J. P. et al.) (Cochrane, 2021).

57. Ioannidis, J. P. & Trikalinos, T. A. An exploratory test for an excess of significant findings. *Clinical Trials* **4**, 245–253 (2007).

58. Papadimitriou, N. *et al.* An umbrella review of the evidence associating diet and cancer risk at 11 anatomical sites. *Nature Communications* **12**, 4579 (2021).

Table 1

Quality assessment for studies providing unique effects

First Author	Year	Quality Assessment: Elig. Crit.[1]	Lit. Search[2]	Dual Screen[3]	Dual Qual.[4]	Studies Listed[5]	Pub. Bias[6]	Hetero.[7]
Abrami	2020	U	U	L	H	L	L	L
Adelantado-Renau	2019	L	L	L	L	L	L	L
Andrade	2019	U	L	L	U	L	H	L
Arztmann	2022	U	H	H	H	H	L	L
Aspiranti	2020	U	L	L	H	L	H	L
Bartel	2015	L	L	U	U	L	U	U
Beck Silva	2022	L	L	L	L	L	H	L
Benavides-Varela	2020	U	H	L	H	L	L	L
Blok	2002	U	L	H	H	L	H	L
Bossen	2020	U	L	L	L	L	H	L
Boyland	2016	H	L	L	U	L	L	L
Byun	2018	U	U	U	H	H	H	H
Cao	2020	U	H	U	H	L	L	L
Champion	2019	L	L	L	L	L	L	L
Chan	2014	U	H	H	H	L	L	L
Chauhan	2017	U	L	U	H	H	L	L
Chen	2020	U	H	U	H	H	H	L
Cheung	2012	U	L	L	H	H	L	L
Cheung	2013	L	H	H	U	L	L	L
Cho	2018	U	H	U	H	L	L	L

Table 1

Quality assessment for studies providing unique effects (continued)

First Author	Year	Elig. Crit.[1]	Lit. Search[2]	Dual Screen[3]	Dual Qual.[4]	Studies Listed[5]	Pub. Bias[6]	Hetero.[7]
Claussen	2022	U	L	U	H	L	H	L
Clinton	2019	U	H	U	U	L	L	L
Comeras-Chueca	2021	L	U	L	U	L	H	L
Comeras-Chueca	2021	L	L	L	U	L	H	L
Coyne	2018	L	L	L	H	L	L	L
Cunningham	2021	U	L	L	H	L	L	L
Cushing	2010	U	L	H	H	L	L	L
Darling	2017	U	L	U	U	L	H	H
Eirich	2022	U	L	L	L	L	L	L
Feng	2021	L	L	L	L	L	H	L
Ferguson	2017	U	L	L	H	L	L	L
Ferguson	2020	L	U	L	L	L	L	L
Folkvord	2018	U	L	L	U	L	H	L
Furenes	2021	H	H	L	U	L	L	L
Gardella	2017	U	L	L	U	L	L	L
Garzón	2019	U	H	U	H	H	L	L
Graham	2015	U	L	H	H	L	L	L
Hammersley	2016	L	L	H	L	L	H	L
Hao	2021	U	L	L	L	L	H	L
Hassan-Saleh	2019	U	L	U	U	H	H	L
He	2021	L	L	L	L	L	L	L

Table 1

Quality assessment for studies providing unique effects (continued)

First Author	Year	Elig. Crit.[1]	Lit. Search[2]	Dual Screen[3]	Dual Qual.[4]	Studies Listed[5]	Pub. Bias[6]	Hetero.[7]
Hernandez-Jimenez	2019	U	L	H	L	L	L	L
Hurwitz	2018	L	L	H	H	L	L	L
Ivie	2020	U	L	L	L	L	L	L
Janssen	2020	U	L	L	L	L	U	L
Kates	2018	U	H	L	H	H	L	L
Kim	2021	U	L	U	L	L	L	L
Kroesbergen	2003	U	L	U	H	L	H	L
Kucukalkan	2019	U	L	U	U	H	L	L
Li	2010	U	L	L	U	L	H	L
Li	2022	L	H	L	L	L	H	L
Li	2022	U	H	L	H	L	L	L
Liao	2008	L	H	H	L	H	H	H
Liao	2014	U	L	H	L	L	L	L
Liu	2019	U	L	U	H	L	L	L
Liu	2022	U	H	U	H	H	L	L
Lu	2021	U	L	U	L	L	L	L
Madigan	2020	U	L	L	U	L	L	L
Major	2021	U	L	L	H	L	L	L
Mallawaarachchi	2022	L	L	L	L	L	L	L
Mares	2005	U	L	H	H	L	H	H
Mares	2013	U	H	H	H	L	H	L

Table 1

Quality assessment for studies providing unique effects (continued)

First Author	Year	Elig. Crit.[1]	Lit. Search[2]	Dual Screen[3]	Dual Qual.[4]	Studies Listed[5]	Pub. Bias[6]	Hetero.[7]
Marker	2022	U	L	H	L	L	L	L
Marshall	2004	U	L	H	H	H	H	L
Martins	2019	U	L	U	H	L	L	L
Martins	2022	L	L	L	L	L	H	L
Mazeas	2022	L	L	L	L	L	L	L
McArthur	2012	L	L	L	L	L	L	L
McArthur	2018	L	L	L	L	L	L	L
Mei	2018	U	H	U	L	L	H	L
Merchant	2014	U	L	H	H	H	H	L
Neitzel	2022	U	L	H	H	L	H	H
Oldrati	2020	U	L	U	H	L	L	L
Paik	1994	U	H	U	H	H	L	H
Pearce	2016	U	L	H	H	H	L	L
Peng	2011	U	L	U	U	L	H	L
Powers	2013	U	L	U	H	L	L	L
Prescott	2018	U	L	U	H	L	L	L
Reynard	2022	H	L	L	L	L	L	L
Rodriguez-Rocha	2019	U	L	L	L	L	L	L
Sadeghirad	2016	H	L	L	L	L	L	L
Scherer	2020	U	H	U	H	L	L	L
Schroeder	2013	L	L	U	H	L	L	L

Table 1

Quality assessment for studies providing unique effects (continued)

First Author	Year	Elig. Crit.[1]	Lit. Search[2]	Dual Screen[3]	Dual Qual.[4]	Studies Listed[5]	Pub. Bias[6]	Hetero.[7]
Scionti	2019	L	L	L	H	L	L	L
Shin	2019	U	L	L	L	L	H	L
Shin	2022	L	H	L	L	L	L	L
Slavin	2014	U	H	H	H	L	H	H
Strouse	2021	U	L	U	H	H	L	L
Takacs	2014	H	L	U	H	L	L	L
Takacs	2019	L	L	U	H	L	L	L
Tekedere	2016	U	H	U	U	L	L	L
Tokac	2019	U	H	L	H	L	L	L
Vahedi	2018	L	L	U	U	L	L	L
van Ekris	2016	U	L	L	L	L	H	L
Vannucci	2020	U	L	U	H	L	L	L
Williams	1982	U	U	H	U	L	H	H
Wouters	2013	U	H	U	H	L	L	L
Xie	2018	U	L	L	H	L	L	L
Yin	2019	U	H	U	H	L	L	L
Zhou	2020	U	L	U	H	L	L	L

Zucker	2009	L	L	U	H	L	H	L

Note: Items are from the National Health, Lung and Blood Institute's Quality Assessment of Systematic Reviews and Meta-Analyses tool. Note that we excluded the first item of the tool. U = Unclear; L = Low; H = High [1] Eligibility criteria predefined and specified [2] Literature search strategy comprehensive and systematic [3] Dual independent screening and review [4] Dual independent quality assessment [5] Included studies listed with important characteristics and results of each [6] Publication bias assessed [7] Heterogeneity assessed