Clinical Psychology Review 67 (2019) 22–35

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

# Diversity in randomized clinical trials of depression: A 36-year review

Antonio J. Polo[a,*], Bridget A. Makol[a,b], Ashley S. Castro[a], Nicole Colón-Quintana[a], Amanda E. Wagstaff[a,c], Sisi Guo[d]

[a] DePaul University, Department of Psychology, 2219 N. Kenmore Ave., Chicago, IL 60614, United States
[b] University of Maryland, College Park, Department of Psychology, 4094 Campus Drive, Room 3123K, College Park, MD 20742, United States
[c] University of California, San Francisco, Department of Psychiatry, Zuckerberg San Francisco General Hospital, 1001 Potrero Ave., Building 5, Suite 6B, United States
[d] University of California, Los Angeles, Semel Institute for Neuroscience and Human Behavior, 760 Westwood Blvd., Los Angeles, CA 90024, United States

## HIGHLIGHTS

- 342 randomized control trials for depression were coded to examine sample diversity.
- Only some participant demographics (e.g., ethnicity) are increasingly being reported.
- Several ethnic groups are not well represented and linguistic minorities are excluded.
- Effects across ethnic groups and ethnicity moderation analyses are rarely reported.
- Increased consideration of diversity in RCTs would improve their generalizability.

## ARTICLE INFO

*Keywords:*
Randomized clinical trials
Diversity
Representation
Depression
Generalizability
Psychotherapy

## ABSTRACT

Historically, authors reporting the results of randomized clinical trials (RCTs) to address mental health problems have insufficiently described sample characteristics pertaining to the ethnic/racial, linguistic, socioeconomic, and immigrant backgrounds of participants. RCTs have also had inadequate representation of participants from diverse backgrounds. This study reports on the trends in the reporting and representation of various sample demographic characteristics in RCTs of psychotherapy and other psychosocial interventions for depression over a 36-year period, and on the extent to which ethnicity, in particular, is considered in the analyses of treatment effects. A total of 342 trials (85.1% comprised of adult samples), representing 61,283 participants, are summarized in the review. Reporting for ethnicity and socioeconomic indicators improved over time, and RCTs for depression have also increasingly included significant numbers of ethnic minority and low-income groups. However, trials are far more likely to exclude, rather than include, linguistic minorities, and have not enrolled a meaningful number of Asian American, Native Hawaiian/Pacific Islander, Native American/Native Alaskan and multi-ethnic participants. Finally, treatment effects are almost never presented separately across racial/ethnic groups and ethnicity moderation analyses are only sporadically conducted. These findings have implications for generalizability, policy, journal reporting guidelines, and dissemination and implementation.

## 1. Introduction

As the United States (U.S.) population has grown, it has also become increasingly diverse. For instance, as of 2017, almost 40% of the U.S. population, or approximately 125 million people, were of ethnic minority backgrounds (U.S. Census Bureau, 2017), and over half of the population is expected to belong to an ethnic minority group by 2044 (Colby & Ortman, 2015). There are over 42 million U.S. residents who are foreign-born, and 85% of them speak a language other than English at home (Colby & Ortman, 2015; Gambino, Acosta, & Grieco, 2014). Furthermore, a substantial proportion of U.S. residents are considered to be disadvantaged socio-economically. The poverty rate in the U.S. currently stands at 12.7%, a rate that has not dramatically improved in the past decade (Semega, Fontenot, & Kollar, 2017). Calls for better representation of diversity in psychological research in the U.S. have long been made (Graham, 1992; Guthrie, 1976; Hall, Yip, & Zárate, 2016; Sue, 1999; Sugden & Moulson, 2015), and research in clinical trials is no exception (Huey & Polo, 2008; Nayayama Hall, 2001). Despite early evidence that ethnicity and socio-economic status (SES) impact psychotherapy treatment effects (Smith & Glass, 1977), progress

* Corresponding author.
*E-mail address:* apolo@depaul.edu (A.J. Polo).

https://doi.org/10.1016/j.cpr.2018.09.004
Received 13 August 2018; Received in revised form 18 September 2018; Accepted 21 September 2018
Available online 22 September 2018

has been slow, difficult to track, and not well documented.

The need for inclusion of women and ethnic minority participants, specifically, was addressed by the NIH Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research (1994), which were put in place in direct response to the passage of the NIH Revitalization Act of 1993 (Public Law 103-43; U.S. Congress, 1993). Among the stipulations included in these guidelines were that NIH would: a) ensure that women and ethnic minorities were included in research; b) ensure that women and ethnic minorities were included sufficiently to allow for analyses of differences in intervention effects (for "Phase III clinical trials"); and c) initiate programs and support for recruitment and outreach efforts of these groups (Hohmann & Parron, 1996). Although it has been 25 years since these guidelines were established, it is not clear to what extent these goals have been achieved.

Publishing standards, which specify journal reporting requirements, also play a key role in the degree to which diversity is prioritized in clinical research studies. In particular, two sets of standards, the 2001 and 2010 Consolidated Standards of Reporting Trials (CONSORT; Moher, Schulz, & Altman, 2001; Moher et al., 2010) and the 2008 Journal Article Reporting Standards (JARS; APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008), have a direct impact on published and peer-reviewed research articles describing RCTs. Both the CONSORT and JARS aim to improve the quality and transparency of reporting of RCT design and characteristics (e.g., eligibility criteria, primary outcome measures). Since its publication in 2001, reporting of several CONSORT items have improved, particularly when journals encouraged or required the use of the CONSORT statement (Turner et al., 2012). In 2010, the statement was updated with several changes, such as increased emphasis on specificity when describing trial design, interventions, and recruitment. Of note, both the CONSORT and JARS specify that descriptions regarding participants must be included, but do not indicate which demographic characteristics must be reported. Thus, even when adhering to CONSORT or JARS guidelines, published RCTs may lack critical information needed for determining whether findings extend to diverse populations.

Examining time trends allows for a closer examination of progress in inclusion and reporting of participants of diverse backgrounds, and can aid in determining whether federal and publication guidelines have had a significant impact. This review compares the reporting practices and levels of representation of key demographic groups across three 12-year time periods, making it possible to compare one period that predates the publication of the NIH guidelines (1981–1992), and two others that include the publication of the CONSORT and JARS (1993–2004 and 2005–2016). Both reporting practices and levels of representation are emphasized, since improvements in representation cannot be tracked without adequate reporting. The focus is specifically on depression trials and on psychosocial interventions and several components of diversity are included, namely the age, gender, ethnicity, SES, immigrant background, and language of participants enrolled in RCTs over a 36-year period.

### 1.1. Representation of ethnic minority participants in RCTs

Efforts to synthesize the literature have consistently made evident the need to improve both the representation of ethnic minorities and reporting of ethnicity in clinical trials. This was strikingly noted in the Surgeon General's landmark report, *Mental Health: Culture, Race, and Ethnicity* (DHHS, 2001). Data for a 15-year period were pooled across RCTs targeting major depression, bipolar disorder, schizophrenia, and Attention Deficit Hyperactivity Disorder (ADHD). Ethnicity was reported for fewer than half of nearly 10,000 participants, and only 671 ethnic minority participants were reported as being included (561 African American, 99 Latino/a, 11 Asian American/Pacific Islander, and 0 American Indian/Alaska Native) across all trials for these conditions (Miranda, Nakamura, & Bernal, 2003).

Mak, Law, Alvidrez, and Pérez-Stable (2007) more specifically reviewed published clinical trials funded by the NIH during a 10-year period (1995–2004) following the passage of the NIH Revitalization Act of 1993. They searched five major mental health journals (four psychiatry and one psychology) and found that over one in four (26.6%) studies did not report any information about the sample's ethnicity and fewer than half (47.8%) provided complete ethnicity breakdowns for their samples. The authors did not provide separate analyses for psychosocial and pharmacological trials or by age group. Similarly, Braslow et al. (2005) reviewed medication and psychotherapy clinical trials published between 1981 and 1996 across 414 studies of predominantly adult samples published in four journals (two psychology and two psychiatry). Although a significant improvement was found in the inclusion of ethnic minority participants over this 15-year period, 71% of studies did not report participant ethnicity and only 25% explicitly reported including participants of ethnic minority backgrounds. Weisz, Jensen Doss, and Hawley (2005) summarized findings of RCTs conducted on youth samples between 1962 and 2002 and focusing on psychotherapy for anxiety, depression, ADHD/impulsivity, and conduct problems. Out of 236 trials, the majority (59.7%) did not report on participants' ethnicity. This review, however, did not separately report on studies conducted inside and outside the U.S. A more recent and comprehensive meta-analysis of studies of the effects of psychotherapy on children and adolescents spanning five decades revealed that only 12.1% of the almost 450 studies featured samples in which the majority of the sample was comprised of ethnic minority participants (Weisz et al., 2017).

Although progress has been noted (Huey, Tilley, Jones, & Smith, 2014), reports of ethnic minority representation in RCTs are in need of more recent data, and limited information is available for trials across both child and adult populations. It is also not known to what extent the representation of ethnic minorities has improved in RCTs for depression that have focused on, or included, psychosocial interventions. Furthermore, previous reviews have too often combined ethnic minority groups to evaluate their representation, thus limiting the utility of the data for any particular ethnic group. Therefore, this review will emphasize, whenever possible, the representation and reporting of each of the major U.S. ethnic groups.

### 1.2. Representation of low income, linguistic minority, and immigrant participants in RCTs

There is limited knowledge on the reporting practices in published RCTs for a number of demographic variables beyond ethnicity, such as income, education, immigrant background, and language proficiency. Crosby et al. (2010) included some of these variables in an analysis of empirical studies across four leading psychology journals between 2004 and 2008. High variability across reporting was found across the 1031 articles coded, ranging from 28.2–59.5% for participant education and from 3.4–29.1% for participant income. Separate analyses were not conducted, however, for RCTs. In their summary of RCTs for youth populations, Weisz et al. (2005) reported that almost three-quarters (72.5%) of studies did not provide any information on the sample's SES.

Although specific clinical trials have been conducted focusing on low-income populations (e.g., Alegría et al., 2014; Miranda, Chung, et al., 2003), individuals from low-income and lower education backgrounds are not well represented in RCTs (Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012). This is problematic as treatment outcomes have been found to be moderated by SES (Falconnier, 2009; Rieppi et al., 2002; Thompson-Brenner et al., 2013). A meta-analysis of school-based interventions among low income, urban youth found negligible effects for these programs, highlighting the potential variability in treatment efficacy when SES is considered (Farahmand, Grant, Polo, & Duffy, 2011). Furthermore, a recent meta-analysis of psychotherapy interventions for depression and anxiety among adult patients found some evidence that socioeconomic deprivation is associated with poorer

treatment response (Finegan, Firth, Wojnarowski, & Delgadillo, 2018).

Even less is known about the patterns of reporting and levels of representation of participants of linguistic minority and immigrant backgrounds in U.S. clinical trials. It appears that these demographic variables are among the least well reported in clinical trials. In Crosby et al.'s (2010) study, 0–10.1% of articles reported on participant language, and 1.0–11.4% reported on participant language proficiency. Of note, one of the four journals included in their review is specifically devoted to the study of culture, race, and ethnicity, and thus likely overestimates the reporting practices in journals, overall. Although there are notable exceptions (e.g., Kataoka et al., 2003; Telles et al., 1995), few RCTs routinely include or focus on immigrant or linguistic minority populations.

### 1.3. A focus on depression

Around the world, depression impacts over 300 million individuals, or roughly 4.4% of the population. In the U.S., the lifetime prevalence of Major Depressive Disorder (MDD) among adults is 16.2% (Kessler et al., 2003), and approximately 17.5 million people are affected by depression, with the highest total and percentage (5.9%) in the Americas (WHO, 2017). The number of years of disability associated with MDD is among the highest of all chronic conditions impacting individuals in the U.S. (Murray et al., 2013). Gender differences begin in adolescence (Merikangas et al., 2010), and it has been consistently found that depression is about twice as prevalent among women compared to men (Piccinelli & Wilkinson, 2000). Additionally, there is substantial evidence that risk for depressive disorders has steadily increased among younger cohorts (Hidaka, 2012).

The prevalence of depressive disorders among U.S. adults varies by ethnicity and nativity. For example, relative to European Americans, African American adults report significantly lower odds of having a depressive disorder (Breslau, Kendler, Su, Gaxiola-Aguilar, & Kessler, 2005). Among children and adolescents, higher risk for depressive symptoms and depressive disorders has consistently been found for some U.S. ethnic minority groups, especially Latino/as (Merikangas et al., 2010; Wagstaff & Polo, 2012). Regardless of ethnicity, U.S.-born adults consistently report significantly higher lifetime risk of depressive disorders than foreign-born adults (Alegría, Canino et al., 2008; Grant et al., 2004; Takeuchi et al., 2007).

Disparities across ethnic and nativity groups are particularly evident in the access and use of services to address depression (Alegría, Chatterji, et al., 2008; González, Tarraf, Whitfield, & Vega, 2010; Merikangas et al., 2011). Among U.S. adults with a past-year depressive disorder, Asian Americans, Latino/as, and African Americans were 48–66% less likely to have received mental health care, relative to European Americans (Alegría, Chatterji, et al., 2008). Similarly, U.S. adolescents with mood disorders who were of ethnic minority backgrounds were 53–77% less likely to receive mental health care, relative to their European American counterparts (Merikangas et al., 2011). Another consistent pattern related to service use is that individuals who are foreign-born and have limited English proficiency are substantially less likely to access mental health care than their counterparts who are born in the U.S. and English-proficient (Abe-Kim et al., 2007; Georgiades, Paksarian, Rudolph, & Merikangas, 2018; Kim et al., 2011).

### 1.4. Research as usual in depression RCTs

Interventions focusing on depression have been extensively researched over the past four decades, with over 500 RCTs conducted (Cuijpers, 2017), allowing for the evaluation of trends across time. However, remarkably little is known from RCTs about the effects of evidence-based depression treatments across diverse groups. This is in large part because, even when samples are diverse, it is not a common practice to report on the efficacy of interventions across subgroups of participants within a trial. In addition, despite their potential to uncover which treatments are best suited for which types of clients, moderation analyses are underutilized in the treatment outcome literature (Kazdin, 2007). A recent review of depression trials suggests that demographic characteristics are not commonly evaluated as either predictors or moderators of treatment effects (Weersing, Jeffreys, Do, Schwartz, & Bolano, 2017). Although investigators conducting meta-analyses have pooled data to determine if factors such as age, gender, and ethnicity are associated with treatment effects (Cuijpers, Karyotaki, Reijnders, & Huibers, 2018), the vast majority of interventions or treatment modalities for depression are not classified as evidence-based for individuals of ethnic minority, low income, immigrant, and linguistic minority backgrounds (Huey & Polo, 2008). Stated otherwise, given that most “evidence” for evidence-based treatments comes from samples of primarily European American and middle-to-high class individuals (Huey & Polo, 2008; Mak et al., 2007), it is unclear whether the effects of these treatments generalize to other groups. This is unfortunate given that our field is currently shifting focus towards the dissemination and implementation of protocols that have been evaluated in clinical trials, and to train and support “usual care” providers in delivering evidence-based care. A number of barriers have been identified in implementing evidence-based protocols, many of which pertain to discrepancies between the settings in which protocols are developed and evaluated and the settings in which they are eventually delivered. Usual care providers report significant concerns about the applicability of established interventions to their clients, who are characterized by a diversity not represented in the literature (Southam-Gerow, Rodríguez, Chorpita, & Daleiden, 2012). Importantly, participants from ethnic minority and low-SES backgrounds have been less likely to access services, less engaged in treatment, and have higher likelihoods of premature termination and reduced benefits from treatment (Curry et al., 2006; Fortuna, Alegría, & Gao, 2010; Interian, Lewis-Fernández, & Dixon, 2013; Miller, Southam-Gerow, & Allin, 2008), which may be due in part to existing interventions not being designed for, or evaluated with, these groups.

These findings illustrate how important it is to consider the role of diversity when implementing depression treatment protocols. As a first step, it is important to understand the extent to which published depression RCTs include individuals from diverse backgrounds. Researchers conducting meta-analyses have lamented that treatment outcome effects cannot be fully evaluated across demographic groups due to lack of reporting and representation (Huey & Polo, 2008; Weersing et al., 2017). Subsequently, this review examines progress in this regard, by describing “research as usual” with respect to the inclusion and representation of diverse samples in clinical trials of depression that incorporate psychosocial interventions.

### 1.5. Rationale and aims

The field needs a more systematic examination of the degree to which RCTs for depression include diverse groups and report on the demographic characteristics of participants. Prior research has surveyed only select journals, and has spanned limited timeframes. It is critical to evaluate the extent to which established policies and guidelines have impacted the representation and reporting of diverse populations in clinical trials in the context of a shifting demographic landscape in the U.S. over the past four decades. Specifically, this study aims:

1) To examine the degree to which published RCTs of psychosocial interventions for depression with children and adults conducted in the U.S. ***report*** the sample's demographic characteristics (gender, age, ethnicity, income, education, language, and immigrant background). Also, to evaluate time trends in the reporting practices from 1981 to 2016.
2) To examine the ***representation*** in published RCTs of psychosocial interventions for depression with children and adults conducted in

**Table 1**
Examples of types of variables coded in the systematic review.

| Variable category | Coding values | Example variable | Example coding guidelines | Demographic variables coded |
|---|---|---|---|---|
| General demographic reporting[a] | 1 = Yes, 0 = No | Income | Are the income levels of the participants reported? Code "Yes" even if grouped into categories (e.g., % above poverty or above a certain income level). | Age, ethnicity, education level, income, gender, immigrant background |
| Reporting of bias analyses[a] | 1 = Yes, 0 = No | Gender | Is there an analysis to evaluate if, after randomization, assigned groups varied across gender? | Age, ethnicity, SES, gender, immigrant background |
| Demographic information extracted or calculated[b] | Continuous, NS = −999 | Age | What is the mean age of the sample of participants? | Age, ethnicity, gender, immigrant background |
| Inclusion of selected populations[a,c] | 1 = Yes, 0 = No | SES | Do the authors report that the sample consisted predominantly of participants that are low income, poor or low SES? | SES, Language |
| Exclusion of selected populations | 1 = Yes, 0 = No, NS = −999 | Language | Were participants (youth or parents) excluded from the study because of limited English proficiency? | Language |
| Moderation analyses | 1 = Yes, 0 = No | Ethnicity | Were moderation analyses conducted to determine if effects varied across ethnic groups? | Ethnicity |

*Note:* NS = not specified; SES = socioeconomic status.
[a] In child trials, SES was coded if parental income or parental education levels were reported; grade was allowed as an indicator of age/developmental level; immigrant background was coded as reported if either child or parental nativity was reported.
[b] For ethnicity, gender, and immigrant background - percentages were extracted or calculated from sample n, if that is all that was reported; in child trials, if only child's grade was reported, age was coded as missing.
[c] For language – used to code studies which explicitly mentioned that participants who spoke a language other than English were included in the study (e.g., Chinese, Spanish).

the U.S. across specific demographic groups (e.g., ethnic and linguistic minorities, low-SES, immigrant backgrounds). To evaluate time trends in levels of representation for these populations from 1981 to 2016.

3) To examine the extent to which ethnicity is considered in the evaluation of ***treatment effects*** in published RCTs for depression conducted in the US. Specifically, whether treatment effects are reported separately by ethnicity and whether ethnicity moderation analyses are conducted, when applicable. To evaluate time trends in reporting of ethnicity treatment effects and ethnicity moderation analyses from 1981 to 2016.

## 2. Method

### 2.1. Inclusion criteria and search strategies

A trial was included if it: a) included a randomized control design; b) evaluated a psychosocial intervention (e.g., psychotherapy) or strategy targeting depression and compared it to another psychosocial intervention, medication, usual care, placebo, or a wait-list condition; c) was published in a peer-reviewed journal between January 1st, 1981 and December 31st, 2016; d) enrolled participants who were identified as being at-risk for depression (e.g., elevated symptoms of depression) or were diagnosed with a depressive disorder; e) included pre- and post-treatment quantitative comparisons across the randomized groups; f) was conducted in the U.S. or its territories (e.g., Puerto Rico); and g) was written in English.

Our search for articles was conducted using the PsycINFO database. The search query included the following terms: (therapy OR psychotherapy OR cognitive behavioral therapy OR interpersonal therapy OR supportive therapy OR problem-solving therapy OR social skills training OR telehealth OR teletherapy OR e-health OR mindfulness OR commitment therapy OR acceptance therapy OR behavioral activation OR web-based) AND (depression OR depressive OR mood OR major depression OR dysthymia OR depressive disorder) AND (randomized OR randomised OR randomized controlled trial). We supplemented this search with references retrieved through several other sources. Firstly, we downloaded a publicly available database of predominantly adult depression trials from http://www.evidencebasedpsychotherapies.org. This database, which has been used in numerous published meta-analyses, contains 1476 references obtained through searches of abstracts from PubMed, PsychINFO, Embase, and the Cochrane Central Register of Controlled Trials, among other sources (see Cuijpers et al., 2018; Cuijpers, van Straten, Warmerdam, & Andersson, 2008 for details). We obtained a second database of youth trials used in a meta-analysis of psychotherapy interventions for various disorders through correspondence with John Weisz. This database has also been used in published meta-analyses of clinical trials and includes 478 child and adolescent RCTs for a range of mental health problems, including depression, anxiety, conduct problems, and ADHD (Weisz et al., 2017) and was obtained from a comprehensive search of PubMed and PsychINFO databases, among other sources. Finally, we used reference trails from selected published meta-analyses of psychotherapy interventions for depression to search for additional published trials that may have been missed through the sources and searches described earlier.

### 2.2. Coding procedures and codebook

We screened and coded articles using a multi-stage coding protocol. The codebook developed for this project focused on trial design and participant characteristics (e.g., inclusion criteria, ethnic background), as well as analytic and sampling procedures (e.g., sampling approach, moderation analyses). The majority of variables (e.g., "Is the sample's age reported?") were coded for whether the information was reported. In other instances, we extracted data directly from the text (e.g., mean age of participants), or derived it from available information (e.g.,

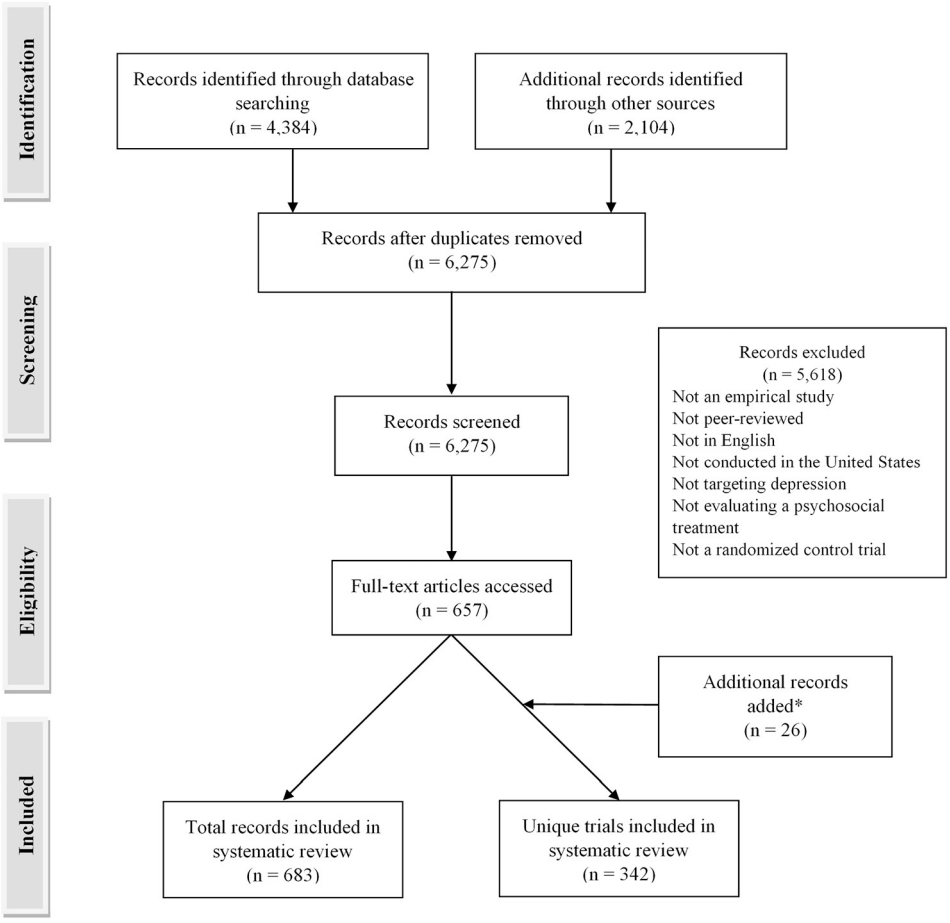


**Fig. 1.** Flowchart of studies included in the systematic review.
*These records were identified through review of previously included records as describing important study design and sample characteristics as well as key outcome analyses.

sample size and mean for each condition). Finally, some variables required inference based on quantitative and qualitative information available in the article. For example, an RCT was coded as focusing on a low-SES population based on information regarding the participant's income and education levels, and other indicators of SES (e.g., unemployment, poverty), in conjunction with the authors' explicit indication that their study enrolled a predominantly low-SES sample. Table 1 includes examples of the types of variables included in the codebook.

Our first stage of coding entailed screening articles for inclusion into the study using the abstract and full-text, when necessary. This initial stage consisted of coding based on the seven eligibility criteria (i.e., a through g) described earlier (see Fig. 1). Screening was conducted by eight coders, who ranged in training from post-baccalaureate to Ph.D.-level, and included all of the study authors. At the second stage of coding, coders extracted information regarding the demographics of participants and other trial design characteristics using the full codebook. We categorized a trial as focusing on adults if the mean sample age was 18 or greater. For child-focused trials, we coded income and education level variables based on reported parent-level demographics. Ethnicity was coded for the seven major U.S. groups. In addition, we included an “ethnic minority” category to accommodate a number of studies in which authors reported the number of non-European American or “ethnic minority” participants but did not further describe their racial or ethnic backgrounds. In other instances, authors provided information about participant ethnicity for only a portion of the sample. We considered a trial as meeting the ethnicity reporting criteria if no more than 5% of the participants were assigned to the “non-specified” category. In both stages of coding, coding discrepancies were resolved by having a third coder investigate, re-code, and consult with the coding team, as necessary.

### 2.3. Search results

As Fig. 1 indicates, our searches and supplemental sources yielded 6488 articles. Of these, 213 were duplicates. The remaining 6275 were screened and a total of 5618 articles were eliminated for one or more of the reasons outlined in Fig. 1 (e.g., no random assignment). For some of these categories (e.g., not conducted in the U.S. or participants not selected for depression/risk), we reviewed the full text of the article to further determine eligibility. The remaining 657 articles were found to be eligible based on the inclusion criteria described above. However, we found that a number of published articles were from the same trial (e.g., one describing the study design and the other reporting the main outcome analyses). In these cases, we collected and coded all identified published articles simultaneously. The “main” study was fully coded and supplemented (as needed) with information from other published articles for the same trial. A total of 26 articles were added to our database to supplement the “main” study articles. We coded 245 trials based on one “main” article, while the remaining 97 trials were coded using data from multiple published articles for the same trial. In total, 342 trials, described across 683 articles, were included. A list of the references for the articles included in this review is available upon request from the first author.

### 2.4. Coding reliability

Inter-rater reliability was examined by double coding articles for both the screening for eligibility and full coding of eligible trials. As noted in prior work (Byrt, Bishop, & Carlin, 1993; Feinstein & Cicchetti, 1990), the kappa statistic, although widely used, is a conservative estimate of reliability that is significantly reduced by low prevalence, even when there is high agreement between coders. In the present study, several variables had low prevalence (i.e., most responses were highly likely to be coded in one category). Thus, we report both the kappa statistic and prevalence-adjusted bias-adjusted kappa (PABAK; Byrt et al., 1993) for inter-rater reliability. We evaluated reliability for categorical variables using the kappa statistic and thresholds described by Viera and Garrett (2005). For continuous variables, we used the Intraclass Correlation Coefficient [*ICC*] statistic and reliability thresholds described by Cicchetti (1994).

For the screening stage, 408 (6.5%) references were double-coded. Coders exhibited substantial agreement in determination of trial inclusion into the next stage of coding ($\kappa = 0.77$, $p < .001$; percentage agreement = 89.5%; PABAK = 0.89). A total of 38 eligible trials (approximately 11%) were also double-coded. Coders exhibited substantial to almost perfect agreement when coding on the reporting of general demographic variables, ($\kappa_{range} = 0.77–1.00$, $\kappa_{mean} = 0.90$, $ps < 0.001$; mean percentage agreement = 97.4%; $PABAK_{range} = 0.89–1.00$, $PABAK_{mean} = 0.95$). They exhibited substantial to almost perfect agreement in coding of inclusion of selected populations ($\kappa_{range} = 0.63–0.83$, $\kappa_{mean} = 0.74$, $ps < 0.001$; mean percentage agreement = 93.4%; $PABAK_{range} = 0.84–0.95$, $PABAK_{mean} = 0.90$) and almost perfect agreement in exclusion of linguistic minorities ($\kappa = 0.87$, $p < .001$; percentage agreement = 97.4%, PABAK = 0.95). Coders exhibited fair agreement in coding of moderation analyses by ethnicity ($\kappa = 0.36$, $p < .05$; percentage agreement = 92.1%; PABAK = 0.84) and moderate agreement in reporting of treatment effects across ethnic groups ($\kappa = 0.48$, $p < .05$; percentage agreement = 94.7%; PABAK = 0.95). For continuous variables, excellent agreement was exhibited for extracted or calculated demographic information ($ICC_{range} = 0.86–1.00$, $ICC_{mean} = 0.96$, $ps < 0.001$).

## 3. Results

As Fig. 1 indicates, a total of 342 studies qualified for inclusion, representing 61,283 participants (adult = 55,418, child = 5,865). Across aims, we grouped trials into three time periods (i.e., A. 1981–1992, B. 1993–2004, C. 2005–2016) to facilitate analysis and interpretation of results. Of the 342 trials identified, more than half ($n = 183$; 53.5%) were conducted in the last time period (period C). Adult trials (85.1%) far outnumbered trials conducted with children/adolescents (14.9%). This trend was evident across all time periods, but particularly in period A, where only 4 out of the 64 trials identified (6.25%) were conducted with children/adolescents. Due to the small sample size for child studies in period A, we limited statistical comparisons across time for child trials conducted in the last two time periods (period B vs. period C).

### 3.1. Aim 1. Reporting of diversity in depression trials

We conducted a series of chi-square analyses to determine whether reporting of participant's demographic characteristics in depression RCTs varied across time. These analyses are presented in Table 2 and separately for adult and child trials.

#### 3.1.1. Race and ethnicity

Overall, 148/342 (43.3%) of the depression RCTs met criteria for reporting on the randomized sample's race/ethnicity. Among adult trials, 43.3% reported on participants' race/ethnicity, and significant differences were found between all three periods. Reporting of race/

**Table 2**
Reporting trends of randomized control trials for depression across key diversity sample demographics: 1981–2016 (N = 342).

| Adult trials | Overall ($n = 291$) | A. 1981–1992 ($n = 60$) | B. 1993–2004 ($n = 76$) | C. 2005–2016 ($n = 155$) | A vs. B $\chi^2$ | A vs. C $\chi^2$ | B vs. C $\chi^2$ |
|---|---|---|---|---|---|---|---|
| Ethnicity | 126 (43.3%) | 10 (16.7%) | 30 (39.5%) | 86 (55.5%) | 8.40** | 26.37*** | 5.23* |
| Education | 212 (72.9%) | 33 (55.0%) | 59 (77.6%) | 120 (77.4%) | 7.85** | 10.60** | 0.001 |
| Income | 86 (29.6%) | 9 (15.0%) | 17 (22.4%) | 60 (38.7%) | 1.18 | 11.16** | 6.13* |
| Immigrant background | 11 (3.8%) | 0 (0.0%) | 2 (2.6%) | 9 (5.8%) | 1.60 | 3.64 | 1.13 |
| Age | 283 (97.3%) | 57 (95.0%) | 74 (97.4%) | 152 (98.1%) | 0.53 | 1.50 | 0.12 |
| Gender | 278 (95.5%) | 54 (90.0%) | 74 (97.4%) | 150 (96.8%) | 3.29 | 4.09* | 0.06 |
| **Child trials** | **Overall ($n = 51$)** | **A. 1981–1992 ($n = 4$)** | **B. 1993–2004 ($n = 19$)** | **C. 2005–2016 ($n = 28$)** | | | **B vs. C $\chi^2$** |
| Ethnicity | 22 (43.1%) | 1 (25.0%) | 8 (42.1%) | 13 (46.4%) | | | 0.09 |
| Education | 22 (43.1%) | 0 (0.0%) | 8 (42.1%) | 14 (50.0%) | | | 0.28 |
| Income | 20 (39.2%) | 0 (0.0%) | 5 (26.3%) | 15 (53.6%) | | | 3.44 |
| Immigrant background | 2 (3.9%) | 0 (0.0%) | 0 (0.0%) | 2 (7.1%) | | | 1.42 |
| Age | 50 (98.0%) | 3 (75.0%) | 19 (100.0%) | 28 (100.0%) | | | n/a |
| Gender | 51 (100.0%) | 4 (100.0%) | 19 (100.0%) | 28 (100.0%) | | | n/a |

*Note.* Income and education represent parental data in child trials. Fisher's exact test is calculated when expected frequencies in each cell are not > 5. Period A for child trials was not used in comparisons due to small sample size.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

ethnicity increased steadily over time such that studies were 3.3 times more likely to report this demographic characteristic in the most recent period, relative to the earliest. Among child trials, 43.1% of RCTs reported on participants' race/ethnicity. There was no significant difference in reporting of child participants' race/ethnicity between periods B and C.

#### 3.1.2. SES

Reporting of SES was coded into two variables: income and education level of participants. Overall, 234/342 (68.4%) of RCTs reported on participants' education and 106/342 (31.0%) reported on their income. Among adult trials, 72.9% reported on participants' level of education and 29.6% reported on their income. Reporting of adult participants' education increased over time across the three time periods, while reporting of adult participants' income increased between periods A and C, and B and C. Among child trials, parental level of education was reported in 43.1% of RCTs, and parental income was reported in 39.2% of RCTs, and no significant differences were found for these variables between periods B and C.

#### 3.1.3. Immigrant background

Overall, 13/342 (3.8%) of depression RCTs reported on participants' immigrant background, with similar levels of reporting in adult (3.8%) and child (3.9%) RCTs. There were no significant differences in reporting of participants' immigrant background across time periods in either adult or child trials.

#### 3.1.4. Age and gender

Overall, the overwhelming majority of depression RCTs reported the age ($n$ = 338; 98.8%) and gender ($n$ = 329; 96.2%) of participants. Among adult trials, 97.3% reported on participants' age, and 95.5% reported on their gender. There were no significant differences in reporting of adult participants' age across time. However, there was a significant increase in reporting of adult participants' gender between periods A and C. Among child trials, 98% reported on participants' age, while 100% reported on their gender. There was no significant difference in reporting of participants' age or gender between periods B and C in child trials.

### 3.2. Aim 2. Diversity representation in depression trials

#### 3.2.1. Low SES

Overall, 44/342 (12.9%) of depression RCTs included a primarily low-SES sample. In the latest time period examined, 18.0% (33/183) of all trials included predominantly low-SES samples (see Table 3). Among adult trials, 12.7% had a primarily low-SES sample, and there were significant increases in rates of trials focusing on this population between periods A and C, and B and C. Among child trials, 13.7% had a primarily low-SES sample, and there was no significant difference between periods B and C in focus on this population.

#### 3.2.2. Ethnic minority/ethnicity

An RCT was coded as focusing on ethnic minority populations when 50% or more of the randomized sample included ethnic minority participants. Overall, 57/342 (16.7%) of depression RCTs met the criteria. In the latest time period examined, almost one in four trials (44/183; 24.0%) had a sample in which a majority of participants were from ethnic minority backgrounds. Among adult trials, 14.4% had a primarily ethnic minority sample, and there were significant increases in rates of trials focusing on these populations between periods A and C, and B and C. Among child trials, 29.4% had a primarily ethnic minority sample, and there was no significant difference between periods B and C in focus on this population.

Fig. 2 shows representation trends for different racial and ethnic groups, combining both child and adult RCTs from 1981 to 2016. European American participants remain the most represented group in

**Table 3**
Focus or inclusion of selected populations in randomized control trials for depression: 1981–2016 (N = 342).

| Adult trials | Overall ($n$ = 291) | A. 1981–1992 ($n$ = 60) | B. 1993–2004 ($n$ = 76) | C. 2005–2016 ($n$ = 155) | A vs. B $\chi^2$ | A vs. C $\chi^2$ | B vs. C $\chi^2$ |
|---|---|---|---|---|---|---|---|
| Low SES (≥50%) | 37 (12.7%) | 2 (3.3%) | 6 (7.9%) | 29 (18.7%) | 1.26 | 8.29** | 4.64* |
| Ethnic minority (≥50%) | 42 (14.4%) | 1 (1.7%) | 7 (9.2%) | 34 (21.9%) | 3.45 | 13.04*** | 5.66* |
| Linguistic minority (any inclusion) | 19 (6.5%) | 1 (1.7%) | 5 (6.6%) | 13 (8.4%) | 1.92 | 3.21 | 0.23 |
| **Child trials** | **Overall ($n$ = 51)** | **A. 1981–1992 ($n$ = 4)** | **B. 1993–2004 ($n$ = 19)** | **C. 2005–2016 ($n$ = 28)** | | | **B vs. C $\chi^2$** |
| Low SES (≥50%) | 7 (13.7%) | 0 (0.0%) | 3 (15.8%) | 4 (14.3%) | | | 0.02 |
| Ethnic minority (≥50%) | 15 (29.4%) | 0 (0.0%) | 5 (26.3%) | 10 (35.7%) | | | 0.46 |
| Linguistic minority (any inclusion) | 4 (7.8%) | 0 (0.0%) | 1 (5.3%) | 3 (10.7%) | | | 0.43 |

*Note.* SES = Socioeconomic status, which includes any indicator such as income or education. Fisher's exact test is calculated when expected frequencies in each cell are not > 5. Period A for Child Trials was not used in comparisons due to small sample size.

* $p < .05$.
** $p < .01$.
*** $p < .001$.

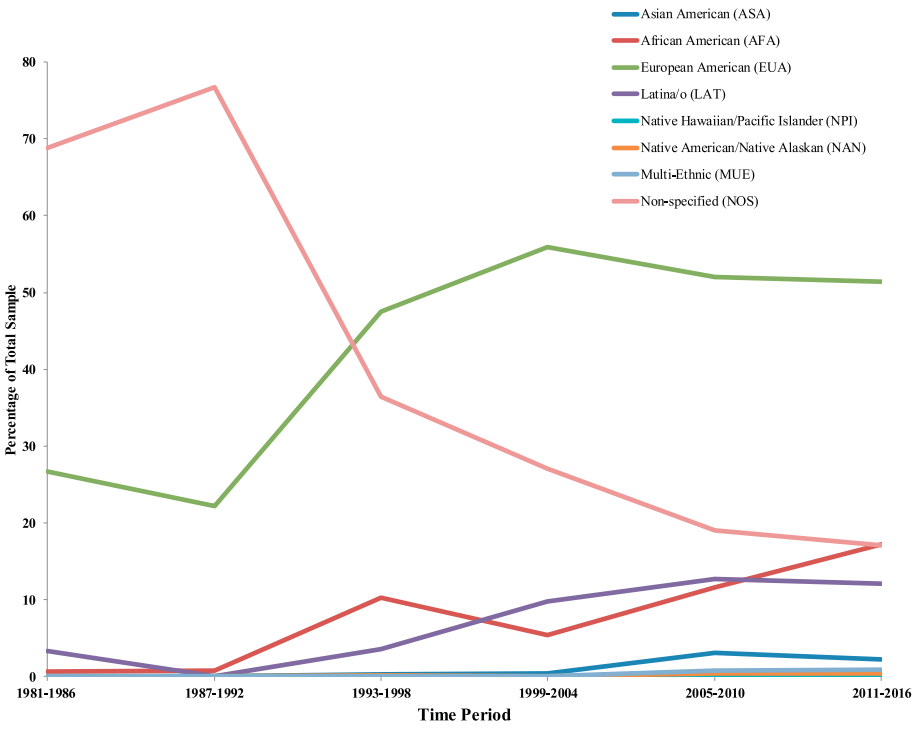


**Fig. 2.** Representation across racial/ethnic groups in child and adult RCTs for depression: 1981–2016 ($N$ = 342); the mean representation across all trials for these groups is: ASA = 1.50%; AFA = 10.06%; EUA = 46.79%; LAT = 8.93%; NPI = 0.06%; NAN = 0.25%; MUE = 0.44%; and NOS = 31.83%.

RCTs over time, making up roughly one half of all participants. African Americans have the next highest representation, but it is almost one fifth that of European Americans. Inclusion of groups such as African Americans and Latino/as has increased over the years and is approaching 20% for the last period displayed for African Americans. Inclusion of Asian American, Native Hawaiian/Pacific Islander, Native American/Native Alaskan, and multi-ethnic participants has remained well below 5% across all time periods. The 'non-specified' group peaked in the period between 1987 and 2002 and has declined since, but remains a sizeable portion even in the latest period examined.

Table 4 further explores the representation of ethnic and racial groups in RCTs for depression. European American and non-specified groups, combined, represent the vast majority of the RCT participants across adult (80.7%) and child trials (67.0%). In adult trials, the mean representation of participants from European American backgrounds (47.2%) was about two and a half times the mean representation of participants from all other ethnic or racial groups combined (19.3%). A similar pattern was observed among child trials, with the exception of Latino/a participants, who had a mean representation of 20.9%. Asian American, Native Hawaiian/Pacific Islander, Native American/Native Alaskan, and multi-ethnic participants had the lowest rates of inclusion across child and adult trials. Broadly, all four groups encompassed 5% or less of the participants in over 90% of the RCTs evaluated in this study. In contrast, the non-specified group represented at least 95% of the sample in over one in four adult trials and in about one in ten child trials.

### 3.2.3. Immigrant background and linguistic minority

The vast majority of trials (96.2%) did not report on the immigrant backgrounds of enrolled participants. Overall, across adult and child trials, we determined that a total of 1514 participants were of immigrant backgrounds, representing 2.5% of all participants between 1981 and 2016.

An RCT was coded as including linguistic minorities if the authors explicitly mentioned any inclusion of participants with limited English proficiency or if the trial was conducted primarily with individuals fluent in languages other than English (e.g., Spanish, Chinese). Overall, 23/342 (6.7%) of depression RCTs reported including linguistic minority participants. Among adult trials, 19/291 (6.5%) included linguistic minority participants (Spanish $n$ = 17, Chinese/Mandarin/Cantonese $n$ = 2), and there were no changes in inclusion rates across time. Among child trials, 7.8% (4/51) included linguistic minority participants (Spanish $n$ = 4), and no significant difference in inclusion was found across time. Overall, based on reporting, we determined that a total of 485 participants were included that spoke a language other than English, representing 0.8% of all participants in depression trials from 1981 to 2016.

We were also interested in examining trends in exclusion of individuals with limited English proficiency from depression trials across time. Overall, 81/342 (23.7%) of depression RCTs reported excluding participants[1] based on English proficiency. Among adult trials, 24.2% reported excluding participants due to limited English proficiency. There was an increase in trials reporting this type of exclusion from periods A (1.7%) to B (22.4%), $X^2$ (2, $N$ = 136) = 14.53, $p$ < .001; and A (1.7%) to C (34.8%), $X^2$ (2, $N$ = 215) = 36.12, $p$ < .001. However, there was no significant difference in reporting of exclusion from periods B (22.4%) to C (34.8%), $X^2$ (2, $N$ = 231) = 5.46, $p$ = .06. Among child trials, 17.6% reported excluding participants based on English proficiency, and there was no significant difference between periods B (15.8%) and C (21.4%), $X^2$ (2, $N$ = 47) = 2.80, $p$ = .25. In the latest period examined, 60/183 (32.8%) of trials reported excluding participants because of limited English proficiency.

### 3.2.4. Gender and age

Based on reporting in trials, we determined that 68.7% ($SD$ = 23.8) of participants in these trials were female (adult = 69.7%,

[1] The vast majority (70.0%) of trials did not report on language inclusion or exclusion criteria for the trial. Among the trials reporting language inclusion or exclusion criteria ($n$ = 81), 71.7% reported excluding those with limited English proficiency.

**Table 4**
Representation of ethnic groups in randomized control trials for depression: 1981–2016 (N = 342).

| Adult trials (n = 291) | ≤ 5.0% | 5.1–19.9% | 20–49.9% | 50–80.0% | 80.1–94.9% | ≥ 95% | Mean representation |
|---|---|---|---|---|---|---|---|
| European American | 102 (35.1%) | 6 (2.1%) | 23 (7.9%) | 68 (23.4%) | 77 (26.5%) | 15 (5.2%) | 47.2% |
| Not Specified | 126 (43.3%) | 57 (19.6%) | 21 (7.2%) | 7 (2.4%) | 3 (1.0%) | 77 (26.5%) | 33.5% |
| African American | 189 (64.9%) | 49 (16.8%) | 36 (12.4%) | 11 (3.8%) | 2 (0.7%) | 4 (1.4%) | 10.2% |
| Latino/a | 213 (73.2%) | 52 (17.9%) | 15 (5.2%) | 4 (1.4%) | 1 (0.3%) | 6 (2.1%) | 6.8% |
| Asian American | 274 (94.2%) | 14 (4.8%) | 1 (0.3%) | 0 (0.0%) | 0 (0.0%) | 2 (0.7%) | 1.5% |
| Multi-ethnic | 285 (97.9%) | 5 (1.7%) | 1 (0.3%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0.3% |
| Native American/Native Alaskan | 287 (98.6%) | 4 (1.4%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0.3% |
| Native Hawaiian/Pacific Islander | 291 (100.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0.1% |
| Child trials (n = 51) | ≤5.0% | 5.1–19.9% | 20–49.9% | 50–80.0% | 80.1–94.9% | ≥ 95% | Mean Representation |
| European American | 18 (35.3%) | 1 (2.0%) | 6 (11.8%) | 13 (25.5%) | 11 (21.6%) | 2 (3.9%) | 44.6% |
| Not specified | 22 (43.1%) | 15 (29.4%) | 6 (11.8%) | 1 (2.0%) | 2 (3.9%) | 5 (9.8%) | 22.4% |
| Latino/a | 26 (51.0%) | 9 (17.6%) | 7 (13.7%) | 4 (7.8%) | 1 (2.0%) | 4 (7.8%) | 20.9% |
| African American | 30 (58.8%) | 15 (29.4%) | 3 (5.9%) | 3 (5.9%) | 0 (0.0%) | 0 (0.0%) | 9.1% |
| Multi-ethnic | 47 (92.1%) | 3 (5.9%) | 1 (2.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1.3% |
| Asian American | 46 (90.2%) | 5 (9.8%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1.2% |
| Native American/Native Alaskan | 50 (98.0%) | 1 (2.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0.2% |
| Native Hawaiian/Pacific Islander | 51 (100.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0.1% |

*Note.* Numbers (and percentages) represent number of trials within each sample representation category. Mean sample representation (last column) refers to the percentage of enrolled participants of each of the ethnic groups among trials which reported ethnicity for at least one of these groups (256 out of 342 trials).

child = 63.7%). On average, adult participants were 47.6 years old ($SD$ = 16.1), with mean age groups including: 19–30 years old (12.1%; $n$ = 39), 31–50 years old (43.5%; $n$ = 141), and 50 years or older (29.1%; $n$ = 94). Child participants were an average of 15.1 years old ($SD$ = 4.47), with mean age groups including: 0–4 years old (0.6%; $n$ = 2), 5–12 years old (2.2%; $n$ = 7), and 13–18 years old (12.4%; $n$ = 40).

### 3.3. Aim 3. Treatment effects across ethnic groups and ethnicity moderation analyses

Beyond representation, we were also interested in the degree to which ethnicity is considered when evaluating treatment effects. We focused on: a) whether authors reported treatment effects separately by ethnicity; and b) whether authors conducted moderation analyses to evaluate differential effects by ethnicity.

#### 3.3.1. Treatment effects across ethnic groups

Overall, only 7/342 (2.0%) depression trials reported treatment effects across ethnic groups. These treatment effects were reported in 2.1% of adult trials, and reporting did not change from periods A (1.7%) to B (2.6%), $X^2$ (1, $N$ = 136) = 0.15, $p$ = .70; B to C (1.9%), $X^2$ (1, $N$ = 231) = 0.12, $p$ = .73; or A to C, $X^2$ (1, $N$ = 215) = 0.02, $p$ = .90. A single child trial (2.0%) reported treatment effects across ethnic groups in period A. Reporting of treatment effects across ethnic groups in child trials did not change between periods B (0%) and C (3.6%), $X^2$ (1, $N$ = 47) = 0.69, $p$ = .41.

#### 3.3.2. Moderation analyses

For these analyses, we restricted the coding to the 135 trials in which participants of at least one racial/ethnic group represented 20–80% of the sample. This variable was coded as "reported" if the trial made explicit mention of conducting moderation analyses of treatment effects by race/ethnicity, regardless of whether the trial reported specific statistics. Overall, 15/135 (11.1%) of depression trials reported ethnicity moderation analyses (12.3% adult trials, $n$ = 13; 6.9% child trials, $n$ = 2). Among adult trials, reporting of ethnicity moderation analyses did not change from periods A (0.0%) to B (14.3%), $X^2$ (1, $N$ = 29) = 0.17, $p$ = .86; B to C (11.7%), $X^2$ (1, $N$ = 105) = 0.13, $p$ = .48; or A to C, $X^2$ (1, $N$ = 78) = 0.13, $p$ = .89. No child trials met the criteria for moderation analysis in period A, and reporting of moderation analyses for child trials did not change from periods B (11.1%) to C (5.0%), $X^2$ (1, $N$ = 29) = 0.36, $p$ = .53. Additionally, we examined reporting of ethnicity moderation analyses for specific racial/ethnic groups. As Table 4 shows, none of the 342 trials included 20–80% of either Native Hawaiians/Pacific Islanders or Native Americans/Native Alaskans. For Asian Americans, only a single adult trial met this criterion, and reported conducting ethnicity moderation analyses. Neither of the two trials (1 child, 1 adult) meeting this criterion for multi-ethnic participants conducted moderation analyses. For the 53 trials meeting this criterion for African American participants, 9/47 (19.1%) adult trials and 0/6 (0%) child trials reported ethnicity moderation analyses. For the 30 trials meeting this criterion for Latino/a participants, 5/19 (26.3%) adult trials and 1/11 (9.1%) child trials reported ethnicity moderation analyses. For the 110 trials meeting this criterion for European American participants, 11/91 (12.1%) adult trials and 2/19 (10.5%) child trials reported ethnicity moderation analyses.

## 4. Discussion

Randomized clinical trials addressing mental health problems have been conducted for many decades and have successfully documented the efficacy of a large number of psychotherapy and psychosocial interventions. Much less is known about how well these treatments fare when delivered across different community settings and to individuals

representing a wide range of socio-demographic and cultural backgrounds. In other words, it is remains unclear for whom, and under which conditions, these interventions work. We conducted a systematic review of depression RCTs carried out in the U.S. over the past four decades to evaluate both the reporting practices and representation of diverse populations in these studies, in order to reflect upon our field's emphasis on documenting the diversity of those who have been enrolled. We tracked reporting of gender and age, but focused particularly on the race/ethnicity, SES, linguistic, and immigrant backgrounds of participants. We separately considered child- and adult-focused trials, and time trends across three 12-year periods from 1981 to 2016 that corresponded with the establishment of key federal agency and journal reporting guidelines. This study summarizes the results based on a total of 61,283 individuals across 342 RCTs. Overall, although the trend is for higher representation of some diverse groups, the findings from this study temper our enthusiasm that the results of hundreds of trials evaluating psychosocial interventions apply broadly to individuals who have not been documented as participants in these treatment evaluations. They suggest that more directed research is needed to engage these participants and ensure their presence.

### 4.1. What are the overall patterns of reporting in RCTs for depression?

Overall, and consistent with previous reports (Crosby et al., 2010; Mak et al., 2007), participant's age and gender are almost always reported by authors in RCTs for depression. Both demographic characteristics were reported in over 95% of the trials. In contrast, fewer than half of the trials fully reported the race/ethnicity of the samples enrolled, and reporting patterns for ethnicity were remarkably similar across both adult and child trials. Mak et al. (2007) also found that, between 1995 and 2004, the vast majority of trials reported on age and gender, while only 47.8% of NIH funded trials in six top journals fully reported the racial/ethnic background of participants.

Reporting of participant education is also more likely than not to be reported, although only among adult trials, where almost three out of every four studies report this participant demographic. Fewer than half of child trials report on the parental levels of education of the individuals enrolled. Participant's education levels are much more commonly reported than participant's income levels, and this discrepancy is especially prominent in adult trials (72.9% vs. 29.6%). Reporting of socioeconomic variables has been notoriously low (Weisz et al., 2005), and our findings indicate that investigators may be more reluctant to ask individuals about their personal or household incomes. Alternatively, it may be that authors are less inclined to report on both indicators when describing the SES of the participants in their articles.

Notably, fewer than 4% of adult and child trials reported on the participant's immigrant background, making it by far the least likely demographic reported of those examined. This is the first study we could identify that systematically tracked the reporting of the nativity of participants enrolled in RCTs. This stands in stark contrast to the literature on the prevalence of psychiatric disorders and access to mental health services where nativity is more routinely considered, and in which findings across U.S.-born and foreign-born individuals are well-established across multiple epidemiological studies (e.g., Merikangas et al., 2011).

### 4.2. What do the time trend analyses tell us about changes in diversity reporting?

As noted earlier, we identified 342 RCTs for depression evaluating psychotherapy and other psychosocial interventions since 1981. Of these, more than half were conducted in the 12-year period between 2005 and 2016. Across time periods, we found that, on average, almost six trials were conducted with adults for every trial conducted with participants under the age of 18. This disparity was largest in the earliest 12-year time period examined (1981–1992), which limited the time trend analyses across the three time periods only to adult trials. Increasing the knowledge base through both prevention and intervention trials with young people is critical given that it is well known that the lifetime rates of depressive disorders rise sharply during adolescence (Merikangas et al., 2010) and that depressive disorders in the U.S. are left untreated an average of four years after onset (Wang et al., 2007).

Improvements in reporting of socio-demographic characteristics are evident in adult trials for ethnicity, income, and education. For ethnicity and income, these trends improved in both of the most recent time periods examined, relative to the first time period. In contrast, none of the time trend comparisons for these variables reached significance for child trials.

### 4.3. Which groups are represented and which groups are missing or excluded?

European Americans, women, and adults ages 31 or older are the most represented participants in RCTs for depression focusing on psychotherapy and other psychosocial interventions conducted over the past four decades. In contrast, relatively few RCTs for depression focus on individuals under the age of 18, and on those of ethnic minority and low socio-economic backgrounds. On average, one of every 10 participants enrolled is under the age of 18. In addition, only about one in six trials (16.7%) enrolled a predominantly ethnic minority sample, and one in seven (12.9%) enrolled a predominantly low SES sample. There is good news in terms of time trends for representation of both of these populations, however. In the most recent period, approximately one in four trials included a predominantly ethnic minority sample (24.0%) and one in 5.5 trials included a predominantly low SES sample.

One of the most disconcerting findings from this review is the number of enrolled participants that have an unspecified ethnic background. Overall, over 1 in 3 adults and over 1 in 5 children who enrolled in RCTs for depression over the past four decades have unreported ethnic backgrounds. Furthermore, there are several groups who are notably absent in this literature, and for whom almost no evidence exists to support the generalizability of RCTs for depression. These include Native Hawaiians/Pacific Islanders, Asian Americans, Native Americans/Native Alaskans, and those of mixed ethnic and racial backgrounds. The mean representation for each of these groups was below 2% and their representation was dismal for both child and adult trials. No single trial included > 5% of Native Hawaiians/Pacific Islanders, only one trial included > 5% of individuals of multiple ethnic backgrounds, and only five trials included > 5% of Native Americans/Native Alaskans. According to U.S. Census data, individuals of two or more races/ethnic groups, along with Asian Americans, are the fastest growing segment of the population (Colby & Ortman, 2015). Additionally, Native Americans have well documented risk in mental health difficulties associated with depression, such as suicidal ideation and suicide (Wexler et al., 2015). Thus, it seems timely to commission studies and provide funding for RCTs conducted specifically with these ethnic groups, to address the current dearth of knowledge in this area.

The two largest ethnic minority groups in the U.S. are Latino/as and African Americans, and they are the two largest ethnic minority groups represented in child and adult RCTs for depression. The overall mean Latino/a representation in RCTs for depression is 8.9%. The mean African American representation in RCTs for depression is 10.1%, and they are roughly equally represented in adult (10.2%) and child (9.1%) trials. Participants of immigrant backgrounds, as well as those who speak a language other than English or who have limited English proficiency are largely absent in published RCTs for depression. Only 13 trials in the 36-year period reported including any participants of immigrant backgrounds and 2.5%, or only one out of every 40 participants, enrolled in the 36-year period were reported as being of immigrant backgrounds. Considering that the percentage of foreign-born individuals in the U.S. stood at 13.3% as of 2016 (Colby & Ortman,

2015), this suggests substantial underrepresentation of individuals of immigrant backgrounds in depression RCTs.

Although 23 trials reported having at least one linguistic minority participant enrolled, fewer than one of every 100 participants enrolled were reported as being linguistic minorities. Although there appears to be an increasing trend in inclusion of participants whose dominant language is not English, none of the time period comparisons revealed statistically significant differences. We found that linguistic minority individuals are not only rarely included in RCTs for depression, but they are increasingly excluded from participation. A total of 58 trials reported excluding non-English speakers, or over 3.5 times more trials than those reporting inclusion of non-English speakers. In the latest time period examined, representing over 50% of all RCTs, one of every three adult trials and one of every five child trials reported limited English proficiency as part of their exclusion criteria. Participants from linguistic minority backgrounds may be increasingly requesting to participate in RCTs, resulting in researchers finding that they are not equipped to enroll them. If true, this is likely due to a lack of translated assessments and treatment protocols, as well as providers who are fluent in languages other than English. Interestingly, Latino/a participants are much less likely to be represented in adult (6.8%) than child (20.9%) trials. It is possible that Latino/a adults may be less likely to be enrolled in adult trials because a substantial number of them are not meeting the investigator's language proficiency requirements.

Improved health care services for individuals with limited English proficiency is part of the mission of the Center for Linguistic and Cultural Competency in Health Care (CLCCHC), as mandated by Congress 30 years ago. Research can support these efforts to increase access and address barriers to care by developing interventions and documenting their impact on linguistic minorities and individuals of immigrant backgrounds (e.g., Kataoka et al., 2003). However, there are unique challenges of conducting rigorous research in the U.S. with participants who are of immigrant backgrounds, as well as with those who are not fluent in English (Hussain-Gambles, Atkin, & Leese, 2004; Lau, Chang, & Okazaki, 2010). Assessment batteries and outcome measures need to be appropriately translated and be psychometrically sound across languages and cultural groups. For clinical trials, the inclusion of linguistic minorities involves translating and/or adapting the assessment and intervention protocol and recruiting and training therapists who can deliver the intervention in a language other than English (Maríñez-Lora, Boustani, del Busto, & Leone, 2016). Similarly, bilingual/multilingual evaluators who can conduct assessments across multiple time points need to be identified and recruited. Investigators without the necessary resources or interest in these populations may not be able or willing to engage in these extra efforts, and thus potentially eligible participants who would benefit from these interventions may be excluded.

### 4.4. How well is ethnicity being considered in the evaluation of treatment effects?

There is an on-going conversation in the treatment outcome literature regarding the importance of tailoring and modifying interventions to meet the needs of individuals of different cultural backgrounds (Nagayama Hall, Ibaraki, Huang, Marti, & Stice, 2016). One important way to inform this debate is to synthesize the current available data on the overall intervention effects across individuals of diverse groups. Part of what has impeded the assessment of treatment effects across studies for individuals of ethnic minority backgrounds, for example, is the limited representation of these groups in RCTs (Weisz et al., 2017).

Our review identified a sizeable number of trials in which ethnicity effects could be evaluated. We found that almost 40% (135/342) of all RCTs included in this 36-year period documented enrolling between 20 and 80% of individuals from at least one ethnic group. Unfortunately, only 2.0% of these trials reported treatment effects across ethnic groups and, on average, only one in nine (11.1%) reported formally evaluating moderation effects by ethnicity. Ethnicity moderation analyses were included in only 15 out of the 83 trials (18.1%) with substantial proportions of either African American or Latino/a participants. Given the particularly low representation of other ethnic minority groups, namely Native Hawaiians/Pacific Islanders, Native Americans/Native Alaskans, and those of mixed ethnic and racial backgrounds, reporting effect sizes separately for these groups or conducting ethnicity moderation analyses is simply not possible. Reporting effect sizes separately for meaningfully represented ethnic groups within a trial is ideal, as is conducting formal moderation analyses to test differential effects, since participants are directly compared when receiving the same intervention protocol in the same setting. Studies with primarily ethnic-minority representation often include culturally-tailored and carefully designed treatments for specific populations. When these studies are compared in meta-analyses to studies in which more generic interventions are evaluated (often with little or no ethnic minority participants), the results can provide limited or misinformed conclusions.

### 4.5. How much impact have NIH policies, the CONSORT statement, and JARS had on the reporting, representation, and evaluation of treatment effects for diverse populations?

A key advantage of looking at the reporting, representation, and evaluation of treatment effects for diverse populations across time periods is to examine if federal policies and journal publication guidelines for RCTs have had a meaningful impact. Since the passing of the Public Health Service Act *sec*. 492B, 42 U.S.C. sec. 289a-2 in 1993, the NIH has been mandated to ensure that women and members of ethnic minority groups in clinical research are included and that clinical trials are designed and conducted in ways that evaluate whether "variables being studied in the trial affect women and ethnic minorities differently than 'others in the trial'." Our systematic review of depression RCTs suggests that reporting of ethnicity and representation of ethnic minority participants increased substantially after 1992. Although the NIH guidelines were not meaningfully revised in the other two periods examined, the trend towards improved reporting of ethnicity and higher ethnic minority representation in these trials continued after 2004. However, there is a lot of room for improvement, suggesting that the goals of these guidelines have not been fully met. First, there are still a large number of trials not reporting ethnicity and not including a meaningful number of ethnic minority participants. Second, and perhaps because the guidelines simply refer to "ethnic minority" participants, representation for several ethnic minority groups has remained extremely low and not improved. Additionally, and equally important, even when investigators are including large and potentially sufficient ethnic minority participants in their trials, they are not investigating if the effects of the interventions are comparable for members of different ethnic groups. NIH guidelines did not include any information about other key demographic groups and these groups are not being included or are being systematically excluded from clinical trials. Specifically, linguistic minorities and individuals of immigrant backgrounds are not being included and there is no evidence that the benefits of psychosocial interventions equally apply to them.

Similarly, CONSORT and JARS publication guidelines have fallen short of providing more specific guidance to authors about ways to include diverse participants and to report findings across demographic groups represented in their samples. Because they are vague in this regard, journals which adopted the original and revised CONSORT and JARS would not have had to encourage that specific demographic characteristics including ethnicity, SES, and linguistic and immigrant background be reported for participants enrolled in RCTs. Moreover, the CONSORT statement, which is specifically geared to improve transparency of the conduct of RCTs and validity of the results, does not recommend to conduct moderation analyses or to report results separately for major demographic groups (Appelbaum et al., 2018). Thus, we recommend that reporting guidelines become more specific, and

journals *require* reporting of key demographic characteristics relevant to understanding the generalizability of RCT effects to diverse populations.

### 4.6. What steps are being taken to increase the validity of RCT trials for diverse groups?

Investigators conducting clinical trials are in some cases required to register their studies in Clinicaltrials.gov, a registry operated by NIH's National Library of Medicine. Sample characteristics that need to be reported include age, sex/gender, and race/ethnicity. The Final Rule for Clinical Trials Registration and Results Information Submission (42 CFR Part 11) has been in effect since January 2017 and requires registration and data submission before and after certain trials are completed (Department of Health and Human Services, National Institutes of Health, 2016), including valid analyses to examine if the study findings extend to women and those of ethnic minority backgrounds. These requirements are currently limited to studies carried out in the U.S. that evaluate at least one drug, biological, or device product regulated by the United States Food and Drug Administration. It remains to be seen whether this enforcement will impact the inclusion and reporting of the trial's sample diversity, and whether trials focused on psychosocial interventions will also be subject to these same standards in the future.

Several journals (e.g., *Child Development*, *Journal of Clinical Child and Adolescent Psychology*) have adopted additional submission and publication requirements that have a direct impact on the reporting and representation of diverse populations. We applaud these efforts, since it is imperative to raise the bar regarding what constitutes sufficient reporting in published RCTs. Given that tremendous energy and resources go into funding, carrying out, and evaluating trials, more strategies are needed to ensure that information about their findings can be more readily obtained. Notably, none of the existing policies or standards encourage researchers or entities to survey the participant's linguistic or immigrant backgrounds, which likely contributes to the paucity of documented research on these populations. Given the dramatic increase in RCTs, as evidenced by the finding that over half of all RCTs for depression were published in the last period examined, it is critical to promptly adopt and enforce improved guidelines that relate to the generalizability and external validity of the findings. Revised guidelines that are connected to requirements from funding agencies and editorial teams are more likely to be embraced by researchers if a rationale is provided regarding the importance and value of diversity not just for social justice and inclusion, but for the advancement of science (Hall et al., 2016).

## 5. Limitations and future directions

Future research in this area could include improved understanding of factors associated with the inclusion of diverse samples in RCTs, and other clinical outcomes studies. For example, it would be helpful to know whether the type of study design (efficacy vs. effectiveness), setting (urban vs. rural; school vs. hospital), source of funding (private vs. public), disorder or condition targeted (substance use vs. anxiety), recruitment or sampling strategy, treatment type (cognitive behavioral vs. medication), or journal characteristics (discipline, impact factor) affect sample diversity. Although factors such as these likely impact whether diverse samples are recruited in clinical research, they are poorly understood.

The findings from this review are specific for RCTs for depression, and may not generalize to RCTs for other mental health conditions. Our hope is that similar examinations and self-reflection regarding diversity can be conducted for other mental health problems. Indeed, some published reviews of clinical trials for other conditions have already reported on inclusion and representation of diverse groups. Mendoza, Williams, Chapman, and Powers (2012) found that fewer than half of RCTs of panic disorder provided ethnicity information. Among those that did, the vast majority of participants were of European American backgrounds (82.7%). Similarly, Strada, Donohue, and Lefforge (2006) found that only about one third of substance abuse RCTs among adolescents fully reported the ethnicity of study participants and European Americans were the majority in two thirds of the 18 studies identified.

This review was limited to studies conducted in the U.S. A large and growing number of studies are being conducted in many other countries. Although it appears that trials conducted outside of the U.S. also lack diversity representation (see Lorenc, Harden, Brunton, & Oakley, 2008), the findings in this report cannot be assumed to be comparable to those from other countries. Similarly, while a good number of the trials included a medication arm or a condition that included a combination of psychotherapy and medication, the review was focused on trials that evaluated the impact of psychosocial interventions so it is not possible to examine diversity consideration in medication RCTs for depression. We also recognize that diversity includes a number of dimensions that were not included in this review. For example, future studies should consider and track reporting and inclusion trends related to other socio-demographic characteristics such as gender identity, sexual orientation, and religious affiliation. In the present study, the findings regarding representation are accurate for studies that reported on the demographic characteristics of the sample. It is possible that this results in an overestimation of diversity representation, to the extent that authors who value diversity more thoroughly inquire about and report participants' sociodemographic and cultural backgrounds. At the same time, diversity may have been underestimated to the extent that individuals of linguistic, immigrant, or multi-ethnic backgrounds were enrolled in RCT trials, but were not included as such because authors did not include questions to more comprehensively assess their backgrounds.

In sum, an opportunity has been missed to expand the scientific knowledge derived from RCTs for depression to a broader and more diverse population. Time trends suggest that, in some ways, our field is beginning to take advantage of this opportunity. The U.S. population's diverse make-up places us in a unique position to provide evidence-based care to members of a wide array of backgrounds and to document and examine factors that determine how well interventions work. For this review, we examined depression, which is a very common and impairing condition. Importantly, concerns have been raised recently that psychotherapy for depression may be less effective now than it used to be (Johnsen Friborg, 2017) and that psychotherapy for depression is less potent than psychotherapy for anxiety and other conditions (Weisz et al., 2017). Evidence shows that accessibility and use of psychotherapy by people suffering from depression is also on the decline (Gaudiano and Miller, 2013) further emphasizing the need to avoid the perception of consumers (clients, parents, and providers) of all backgrounds that this form of treatment is less viable or relevant.

## Author note

The authors wish to thank Alyssa Donovan, Jessica Perez, Nataly Ibarra, Michael Ryva, and the Culture and Evidence-Based Practice Lab members for their enthusiasm, commitment, and dedication to this project. Their help was instrumental in identifying, organizing, and coding the articles presented in this manuscript.

## Declaration of interest

None.

## References

Abe-Kim, J., Takeuchi, D. T., Hong, S., Zane, N., Sue, S., Spencer, M. S., ... Alegría, M. (2007). Use of mental health–related services among immigrant and US-born Asian Americans: Results from the National Latino and Asian American study. *American Journal of Public Health, 97*(1), 91–98. https://doi.org/10.2105/AJPH.2006.098541.

Alegría, M., Canino, G., Shrout, P. E., Woo, M., Duan, N., Vila, D., ... Meng, X. L. (2008). Prevalence of mental illness in immigrant and non-immigrant US Latino groups. *American Journal of Psychiatry, 165*(3), 359–369. https://doi.org/10.1176/appi.ajp.2007.07040704.

Alegría, M., Chatterji, P., Wells, K., Cao, Z., Chen, C. N., Takeuchi, D., ... Meng, X. L. (2008). Disparity in depression treatment among racial and ethnic minority populations in the United States. *Psychiatric Services, 59*(11), 1264–1272. https://doi.org/10.1176/appi.ps.59.11.1264.

Alegría, M., Ludman, E., Kafali, N., Lapatin, S., Vila, D., Shrout, P. E., ... Bauer, A. (2014). Effectiveness of the engagement and counseling for Latinos (ECLA) intervention in low-income Latinos. *Medical Care, 52*(11), 989–997. https://doi.org/10.1097/MLR.0000000000000232.

APA Publications and Communications Board Working Group on Journal Article Reporting Standards (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist, 63*, 839–851. https://doi.org/10.1037/0003-066X.63.9.839.

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal Article Reporting Standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191.

Braslow, J. T., Duan, N., Starks, S. L., Polo, A., Bromley, E., & Wells, K. B. (2005). Generalizability of studies on mental health treatment and outcomes, 1981 to 1996. *Psychiatric Services, 56*(10), 1261–1268. https://doi.org/10.1176/appi.ps.56.10.1261.

Breslau, J., Kendler, K. S., Su, M., Gaxiola-Aguilar, S., & Kessler, R. C. (2005). Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States. *Psychological Medicine, 35*(3), 317–327. https://doi.org/10.1017/S0033291704003514.

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology, 46*(5), 423–429. https://doi.org/10.1016/0895-4356(93)90018-V.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284–290. https://doi.org/10.1037/1040-3590.6.4.284.

Colby, S. L., & Ortman, J. M. (2015). *Projections of the size and composition of the U.S. population: 2014 to 2060, current population reports, P25-1143.* Washington, DC: U.S. Census Bureau2014.

Crosby, J. M., Gundy, J. M., Armstrong, A. B., Nye, E. W., Boman, A., Nelson, C. R., & Twohig, M. P. (2010). How well are we doing at reporting participant characteristics in our research? *The Behavior Therapist, 33*(7), 133–135.

Cuijpers, P. (2017). Four decades of outcome research on psychotherapies for adult depression: An overview of a series of meta-analyses. *Canadian Psychology, 58*(1), 7–19. https://doi.org/10.1037/cap0000096.

Cuijpers, P., Karyotaki, E., Reijnders, M., & Huibers, M. J. (2018). Who benefits from psychotherapies for adult depression? A meta-analytic update of the evidence. *Cognitive Behaviour Therapy, 47*(2), 91–106. https://doi.org/10.1080/16506073.2017.1420098.

Cuijpers, P., van Straten, A., Warmerdam, L., & Andersson, G. (2008). Psychological treatment of depression: A meta-analytic database of randomized studies. *BMC Psychiatry, 8*(1), 1–6. https://doi.org/10.1186/1471-244X-8-36.

Curry, J., Rohde, P., Simons, A., Silva, S., Vitiello, B., Kratochvil, C., ... Weller, E. (2006). Predictors and moderators of acute outcome in the Treatment for Adolescents with Depression Study (TADS). *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(12), 1427–1439. https://doi.org/10.1097/01.chi.0000240838.78984.e2.

Department of Health and Human Services, National Institutes of Health (2016). Final rule for clinical trials registration and results information submission (42 CFR Part 11). Retrieved from: https://www.gpo.gov/fdsys/pkg/FR-2016-09-21/pdf/2016-22129.pdf.

Falconnier, L. (2009). Socioeconomic status in the treatment of depression. *American Journal of Orthopsychiatry, 79*(2), 148–158. https://doi.org/10.1037/a0015469.

Farahmand, F. K., Grant, K. E., Polo, A. J., & Duffy, S. N. (2011). School-based mental health and behavioral programs for low-income, urban youth: A systematic and meta-analytic review. *Clinical Psychology: Science and Practice, 18*(4), 372–390. https://doi.org/10.1111/j.1468-2850.2011.01265.x.

Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I. the problems of two paradoxes. *Journal of Clinical Epidemiology, 43*(6), 543–549. https://doi.org/10.1016/0895-4356(90)90158-L.

Finegan, M., Firth, N., Wojnarowski, C., & Delgadillo, J. (2018). Associations between socioeconomic status and psychological therapy outcomes: A systematic review and meta-analysis. *Depression and Anxiety, 35*(6), 560–573. https://doi.org/10.1002/da.22765.

Fortuna, L. R., Alegría, M., & Gao, S. (2010). Retention in depression treatment among ethnic and racial minority groups in the United States. *Depression and Anxiety, 27*(5), 485–494. https://doi.org/10.1002/da.20685.

Gambino, C. P., Acosta, Y. D., & Grieco, E. M. (2014). *English-speaking ability of the foreign-born population in the United States: 2012. American community survey reports, ACS-26.* Washington, DC: U.S. Census Bureau.

Gaudiano, B. A., & Miller, I. W. (2013). The evidence-based practice of psychotherapy: Facing the challenges that lie ahead. *Clinical Psychology Review, 33*(7), 813–824. https://doi.org/10.1016/j.cpr.2013.04.004.

Georgiades, K., Paksarian, D., Rudolph, K. E., & Merikangas, K. R. (2018). Prevalence of mental disorder and service use by immigrant generation and race/ethnicity among U.S. adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 57*(4), 280–287. https://doi.org/10.1016/j.jaac.2018.01.020.

González, H. M., Tarraf, W., Whitfield, K. E., & Vega, W. A. (2010). The epidemiology of major depression and ethnicity in the United States. *Journal of Psychiatric Research, 44*(15), 1043–1051. https://doi.org/10.1016/j.jpsychires.2010.03.017.

Graham, S. (1992). "Most of the subjects were White and middle class": Trends in published research on African Americans in selected APA journals, 1970–1989. *American Psychologist, 47*(5), 629–639. https://doi.org/10.1037/0003-066X.47.5.629.

Grant, B. F., Stinson, F. S., Hasin, D. S., Dawson, D. A., Chou, S. P., & Anderson, K. (2004). Immigration and lifetime prevalence of DSM-IV psychiatric disorders among Mexican Americans and Non-Hispanic Whites in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry, 61*(12), 1226–1233. https://doi.org/10.1001/archpsyc.61.12.1226.

Guthrie, R. V. (1976). *Even the rat was white: A historical view of psychology.* Pearson Education.

Hall, G. C. N., Yip, T., & Zárate, M. A. (2016). On becoming multicultural in a monocultural research world: A conceptual approach to studying ethnocultural diversity. *American Psychologist, 71*(1), 40–51 (doi: 10.1037%2Fa0039734).

Hidaka, B. H. (2012). Depression as a disease of modernity: Explanations for increasing prevalence. *Journal of Affective Disorders, 140*(3), 205–214. https://doi.org/10.1016/j.jad.2011.12.036.

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36*(5), 427–440. https://doi.org/10.1007/s10608-012-9476-1.

Hohmann, A. A., & Parron, D. L. (1996). How the new NIH guidelines on inclusion of women and minorities apply: Efficacy trials, effectiveness trials, and validity. *Journal of Consulting and Clinical Psychology, 64*(5), 851–855. https://doi.org/10.1037/0022-006X.64.5.851.

Huey, S. J., Jr., & Polo, A. J. (2008). Evidence-based psychosocial treatments for ethnic minority youth. *Journal of Clinical Child & Adolescent Psychology, 37*(1), 262–301. https://doi.org/10.1080/15374410701820174.

Huey, S. J., Tilley, J. L., Jones, E. O., & Smith, C. A. (2014). The contribution of cultural competence to evidence-based care for ethnically diverse populations. *Annual Review of Clinical Psychology, 10*, 305–338. https://doi.org/10.1146/annurev-clinpsy-032813-153729.

Hussain-Gambles, M., Atkin, K., & Leese, B. (2004). Why ethnic minority groups are under-represented in clinical trials: A review of the literature. *Health and Social Care in the Community, 12*(5), 382–388. https://doi.org/10.1111/j.1365-2524.2004.00507.x.

Interian, A., Lewis-Fernández, R., & Dixon, L. B. (2013). Improving treatment engagement of underserved US racial-ethnic groups: A review of recent interventions. *Psychiatric Services, 64*(3), 212–222. https://doi.org/10.1176/appi.ps.201100136.

Johnsen, T. J., & Friborg, O. (2015). The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: A meta-analysis. *Psychological Bulletin, 141*(4), 747. https://doi.org/10.1037/bul0000015.

Kataoka, S. H., Stein, B. D., Jaycox, L. H., Wong, M., Escudero, P., Tu, W., ... Fink, A. (2003). A school-based mental health program for traumatized Latino immigrant children. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*(3), 311–318. https://doi.org/10.1097/00004583-200303000-00011.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1–27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432.

Kim, G., Loi, C. X. A., Chiriboga, D. A., Jang, Y., Parmelee, P., & Allen, R. S. (2011). Limited English proficiency as a barrier to mental health service use: A study of Latino and Asian immigrants with psychiatric disorders. *Journal of Psychiatric Research, 45*(1), 104–110. https://doi.org/10.1001/jama.289.23.3095.

Lau, A. S., Chang, D. F., & Okazaki, S. (2010). Methodological challenges in treatment outcome research with ethnic minorities. *Cultural Diversity and Ethnic Minority Psychology, 16*(4), 573–580. https://doi.org/10.1037/a0021371.

Lorenc, T., Harden, A., Brunton, G., & Oakley, A. (2008). *Including diverse groups of children and young people in health promotion and public health research: A review of methodology and practice.* London: EPPICentre, Social Science Research Unit, Institute of Education, University of London.

Mak, W. W., Law, R. W., Alvidrez, J., & Pérez-Stable, E. J. (2007). Gender and ethnic diversity in NIMH-funded clinical trials: Review of a decade of published research. *Administration and Policy in Mental Health and Mental Health Services Research, 34*(6), 497–503. https://doi.org/10.1007/s10488-007-0133-z.

Maríñez-Lora, A. M., Boustani, M., del Busto, C. T., & Leone, C. (2016). A framework for translating an evidence-based intervention from English to Spanish. *Hispanic Journal of Behavioral Sciences, 38*(1), 117–133. https://doi.org/10.1177/0739986315612769.

Mendoza, D. B., Williams, M. T., Chapman, L. K., & Powers, M. (2012). Minority inclusion in randomized clinical trials of panic disorder. *Journal of Anxiety Disorders, 26*(5), 574–582. https://doi.org/10.1016/j.janxdis.2012.02.011.

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., ... Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry, 49*(10), 980–989. https://doi.org/10.1016/j.jaac.2010.05.017.

Merikangas, K. R., He, J. P., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., ... Olfson, M. (2011). Service utilization for lifetime mental disorders in US adolescents: Results of the National Comorbidity Survey–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry, 50*(1), 32–45. https://doi.org/10.1016/j.jaac.2010.10.006.

Miller, L. M., Southam-Gerow, M. A., & Allin, R. B. (2008). Who stays in treatment? Child and family predictors of youth client retention in a public mental health agency. *Child & Youth Care Forum, 37*, 153–170. https://doi.org/10.1007/s10566-008-9058-2.

Miranda, J., Chung, J. Y., Green, B. L., Krupnick, J., Siddique, J., Revicki, D. A., & Belin, T. (2003). Treating depression in predominantly low-income young minority women: A randomized controlled trial. *JAMA, 290*(1), 57–65. https://doi.org/10.1001/jama.

290.1.57.
Miranda, J., Nakamura, R., & Bernal, G. (2003). Including ethnic minorities in mental health intervention research: A practical approach to a long-standing problem. *Culture, Medicine and Psychiatry, 27*(4), 467–486. https://doi.org/10.1023/B:MEDI.0000005484.26741.79.
Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., ... Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ, 340*, 1–28. https://doi.org/10.1016/j.ijsu.2011.10.001.
Moher, D., Schulz, K. F., & Altman, D. G. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Medical Research Methodology, 1*(1), 1–7. https://doi.org/10.1186/1471-2288-1-2.
Murray, C. J., Abraham, J., Ali, M. K., Alvarado, M., Atkinson, C., Baddour, L. M., ... Bolliger, I. (2013). The state of US health, 1990–2010: Burden of diseases, injuries, and risk factors. *JAMA, 310*(6), 591–606. https://doi.org/10.1001/jama.2018.0158.
Nagayama Hall, G. C., Ibaraki, A. Y., Huang, E. R., Marti, C. N., & Stice, E. (2016). A meta-analysis of cultural adaptations of psychological interventions. *Behavior Therapy, 47*(6), 993–1014.
Nayayama Hall, G. C. (2001). Psychotherapy research with ethnic minorities: Empirical, ethical, and conceptual issues. *Journal of Consulting and Clinical Psychology, 69*(3), 502–510. https://doi.org/10.1037/0022-006X.69.3.502.
NIH Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research (1994). *Federal Register, Part VIII (59 FR 14508–14513), March 28.* (Also reproduced in NIH Guide to Grants and Contracts (Vol. 23, No. 11, March 18, 1994)).
Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression: Critical review. *The British Journal of Psychiatry, 177*(6), 486–492. https://doi.org/10.1192/bjp.177.6.486.
Rieppi, R., Greenhill, L. L., Ford, R. E., Chuang, S., Wu, M., Davies, M., ... Hechtman, L. (2002). Socioeconomic status as a moderator of ADHD treatment outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(3), 269–277. https://doi.org/10.1097/00004583-200203000-00006.
Semega, J. L., Fontenot, K. R., & Kollar, M. A. (2017). *U.S. Census Bureau, Current population reports, p60–259, income and poverty in the United States: 2016.* Washington, DC: U.S. Government Printing Office.
Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*(9), 752–760. https://doi.org/10.1037/0003-066X.32.9.752.
Southam-Gerow, M. A., Rodríguez, A., Chorpita, B. F., & Daleiden, E. L. (2012). Dissemination and implementation of evidence based treatments for youth: Challenges and recommendations. *Professional Psychology: Research and Practice, 43*(5), 527–534. https://doi.org/10.1037/a0029101.
Strada, M. J., Donohue, B., & Lefforge, N. L. (2006). Examination of ethnicity in controlled treatment outcome studies involving adolescent substance abusers: A comprehensive literature review. *Psychology of Addictive Behaviors, 20*(1), 11–27. https://doi.org/10.1037/0893-164X.20.1.11.
Sue, S. (1999). Science, ethnicity, and bias: Where have we gone wrong? *American Psychologist, 54*(12), 1070–1077. https://doi.org/10.1037/0003-066X.54.12.1070.
Sugden, N. A., & Moulson, M. C. (2015). Recruitment strategies should not be randomly selected: Empirically improving recruitment success and diversity in developmental psychology research. *Frontiers in Psychology, 6*, 523. https://doi.org/10.3389/fpsyg.2015.00523.
Takeuchi, D. T., Zane, N., Hong, S., Chae, D. H., Gong, F., Gee, G. C., ... Alegría, M. (2007). Immigration-related factors and mental disorders among Asian Americans. *American Journal of Public Health, 97*(1), 84–90. https://doi.org/10.2105/AJPH.2006.088401.
Telles, C., Karno, M., Mintz, J., Paz, G., Arias, M., Tucker, D., & Lopez, S. (1995). Immigrant families coping with schizophrenia behavioural family intervention v. case management with a low-income Spanish-speaking population. *The British Journal of Psychiatry, 167*(4), 473–479. https://doi.org/10.1192/bjp.167.4.473.
Thompson-Brenner, H., Franko, D. L., Thompson, D. R., Grilo, C. M., Boisseau, C. L., Roehrig, J. P., ... Devlin, M. J. (2013). Race/ethnicity, education, and treatment parameters as moderators and predictors of outcome in binge eating disorder. *Journal of Consulting and Clinical Psychology, 81*(4), 710–721. https://doi.org/10.1037/a0032946.
Turner, L., Shamseer, L., Altman, D. G., Weeks, L., Peters, J., Kober, T., ... Moher, D. (2012). *Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane database of systematic reviews MR000030, 11, 1–161.* https://doi.org/10.1002/14651858 (MR000030.pub2).
U.S. Census Bureau (2017). Quick facts: United States. Retrieved from: https://www.census.gov/quickfacts/fact/table/US/PST045216.
U.S. Congress (1993). *National Institutes of Health Revitalization Act of 1993: Act to Amend the Public Health Service Act to revise and extend the programs of the National Institutes of Health, and for other purposes. Public Law 103-43.* Washington, D.C.: US Congress.
U.S. Department of Health and Human Services (2001). *Mental health: culture, race, and ethnicity: a supplement to mental health: A report of the surgeon general—executive summary.* Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Office of the Surgeon General.
Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine, 37*(5), 360–363.
Wagstaff, A. E., & Polo, A. J. (2012). Ethnicity and adolescent depression: Prevalence, access to services, and promising interventions. *The Prevention Researcher, 19*(4), 8–11.
Wang, P. S., Angermeyer, M., Borges, G., Bruffaerts, R., Chiu, W. T., De Girolamo, G., ... Kessler, R. C. (2007). Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry, 6*(3), 177–185.
Weersing, V. R., Jeffreys, M., Do, M. C. T., Schwartz, K. T., & Bolano, C. (2017). Evidence base update of psychosocial treatments for child and adolescent depression. *Journal of Clinical Child & Adolescent Psychology, 46*(1), 11–43. https://doi.org/10.1080/15374416.2016.1220310.
Weisz, J. R., Jensen Doss, A., & Hawley, K. M. (2005). Youth psychotherapy outcome research: A review and critique of the evidence base. *Annual Review of Psychology, 56*, 337–363. https://doi.org/10.1146/annurev.psych.55.090902.141449.
Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., ... Weersing, V. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *American Psychologist, 72*(2), 79–117. https://doi.org/10.1037/a0040360.
Wexler, L., Chandler, M., Gone, J. P., Cwik, M., Kirmayer, L. J., Lafromboise, T., & Allen, J. (2015). Advancing suicide prevention research with rural American Indian and Alaska Native populations. *American Journal of Public Health, 105*(5), 891–899.
World Health Organization (2017). *Depression and Other Common Mental Disorders: Global Health Estimates. Rep. CC BY-NC-SA 3.0 IGO.* Geneva: World Health Organization.

**Antonio J. Polo** is an Associate Professor in the department of psychology at DePaul University. His research interests concentrate on three primary areas. The first area focuses on the prevention and treatment of depression and anxiety in youth, centering on developing interventions that are evidence-based, transportable, and effectively implemented in community settings, as well as culturally-informed and well-received by ethnic and linguistic minority youth and their families. A second research area involves expanding our understanding about the mental health of Latino children, adolescents, and their families. Specifically, how cultural influences and other risk and protective factors help explain the patterns and levels of maladjustment they exhibit. Finally, Dr. Polo has been involved in developing educational strategies and programs for individuals to become actively involved in the educational and mental health systems. This research includes interventions to help clients become collaborative partners with their providers order to identify and receive the best possible services.