 Original Investigation

School Intervention to Improve Mental Health of Students in Santiago, Chile
A Randomized Clinical Trial

Ricardo Araya, PhD; Rosemarie Fritsch, MD; Melissa Spears, MSc; Graciela Rojas, Dr Med; Vania Martinez, Dr Med; Sergio Barroilhet, PhD; Paul Vöhringer, MD; David Gunnell, PhD; Paul Stallard, PhD; Viviana Guajardo, MD; Jorge Gaete, PhD; Sian Noble, PhD; Alan A. Montgomery, PhD

IMPORTANCE Depression can have devastating effects unless prevented or treated early and effectively. Schools offer an excellent opportunity to intervene with adolescents presenting emotional problems. There are very few universal school-based depression interventions conducted in low- and middle-income countries.

OBJECTIVE To assess the effectiveness of a school-based, universal psychological intervention to reduce depressive symptoms among adolescents from low-income families.

DESIGN, SETTING, AND PARTICIPANTS A 2-arm, parallel, cluster, randomized clinical trial was conducted in secondary schools in deprived socioeconomic areas of Santiago, Chile. Almost all students registered in the selected schools consented to take part in the study. A total of 2512 secondary school students from 22 schools and 66 classes participated.

INTERVENTIONS Students in the intervention arm attended 11 one-hour weekly and 2 booster classroom sessions of an intervention based on cognitive-behavioral models. The intervention was delivered by trained nonspecialists. Schools in the control arm received the standard school curriculum.

MAIN OUTCOMES AND MEASURES Scores on the self-administered Beck Depression Inventory–II at 3 months (primary) and 12 months (secondary) after completing the intervention.

RESULTS There were 1291 participants in the control arm and 1221 in the intervention arm. Primary outcome data were available for 82.1% of the participants. There was no evidence of any clinically important difference in mean depression scores between the groups (adjusted difference in mean, −0.19; 95% CI, −1.22 to 0.84) or for any of the other outcomes 3 months after completion of the intervention. No significant differences were found in any of the outcomes at 12 months.

CONCLUSIONS AND RELEVANCE A well-designed and implemented school-based intervention did not reduce depressive symptoms among socioeconomically deprived adolescents in Santiago, Chile. There is growing evidence that universal school interventions may not be sufficiently effective to reduce or prevent depressive symptoms.

TRIAL REGISTRATION isrctn.org Identifier: ISRCTN19466209

Published online September 2, 2013.
Depression is a common and disabling condition affecting people of all ages, with huge economic consequences. Depression can have devastating life-long effects when onset is early in life. It is common among adolescents in low- and middle-income countries who rarely receive treatment. Depression in adolescence needs to be prevented and, if that is not possible, treated early and effectively.

Psychological interventions, the first-line treatment for depression in adolescence, have proven to be efficacious in clinical settings. However, their efficacy in nonclinical settings remains unproven, especially in low- and middle-income countries. Most low- and middle-income countries do not have enough people trained to deliver treatment or preventive interventions for depressive disorders. Delivering interventions by less-qualified workers, termed task sharing, can help overcome this problem. However, it is unknown whether these workers are capable of delivering effective interventions to reduce depressive symptoms among adolescents in these settings. Schools offer an excellent opportunity to identify adolescents in need, before their emotional problems become entrenched. Several studies have investigated the impact of school-based depression interventions, and virtually all such studies have been conducted in high-income countries.

School-based interventions are classified as universal or targeted: universal interventions cover the entire population at risk, and targeted interventions are conducted only for individuals with a known risk factor (selective) or subthreshold symptoms (indicated). Targeted depression interventions tend to show larger effect sizes than universal interventions. This may be related to the increased severity of symptoms of participants in targeted interventions, or a “floor effect” (unchanged “normals”) affecting universal interventions more markedly. Targeted interventions nonetheless present problems with recruitment and retention and often exclude subsyndromal groups that may benefit. Overall effect sizes with universal interventions have been small, and the quality of many of these studies has been criticized.

To our knowledge, no large randomized clinical trial of a universal, school-based intervention for depressive symptoms from a low- or middle-income country has been published. There is an urgent need to develop potentially scalable programs for adolescents with emotional symptoms in resource-poor settings. Chile is a middle-income country with marked income inequalities. Depression is common among adolescents in Chile, reaching 10% in community surveys. The present randomized clinical trial aimed to test the effectiveness of a universal school-based intervention delivered by nonspecialists to reduce depressive symptoms among low-income secondary school students in Santiago, Chile.

Methods

Study Design
A cluster randomized clinical trial was conducted, with schools as the unit of allocation and individuals as the unit of analysis. Schools were randomized because there was a considerable risk of contamination across classes within the same school. The study was reviewed and approved by the ethics committee of the Clinical Hospital, Faculty of Medicine, Universidad de Chile (No. 179; June 30, 2008).

Setting, Participants, and Eligibility Criteria
Municipal schools provide education for most low-income students in Santiago. Our sampling frame comprised all municipal secondary, mixed-sex schools, with 2 or more 1st Medio classes (equivalent to ninth grade in the United States) in Santiago. Twenty-two schools were selected using stratified random sampling. All students attending 1st Medio grade in the selected schools were eligible and invited to participate, and those with severe depressive episodes and/or clear suicidal risk were encouraged to seek professional advice. Parents were informed of the intervention and advised that they could request the withdrawal of their children from study assessments. Students were asked to sign a written consent form.

Randomization
Randomization took place after the baseline assessment to allow balancing of the study arms with respect to the number of classes in that grade, area of social deprivation, and location of schools. This was achieved by calculating an imbalance statistic for all possible allocation sequences. The trial statistician (A.A.M.) used a computer-generated list of random numbers to select 1 allocation sequence from the 1000 sequences with the most desirable balance properties. Six schools selected had more than 4 eligible classes, and so 4 classes within these schools were randomly selected to participate.

Intervention
The intervention I (Yo), Think (Piensas), Feel (Sientes), and Act (Actas) was based on a cognitive-behavioral therapy model and delivered to the whole class during regular school hours. The intervention was developed after 18 months of formative research. The choice of intervention components was informed by previous school-based depression programs. The intervention consisted of 11 weekly and 2 booster sessions, each lasting approximately 1 hour, similar to other school-based mental health interventions. There was an introductory session, 6 sessions dealing with thought restructuring and emotions, 3 sessions of problem-solving strategies, and 1 closing session to revise and integrate all previous work. Two booster sessions delivered at 2 and 7 months reviewed challenging negative thoughts and problem-solving strategies. A third planned booster session was canceled because 40% of the sample had moved to other classes or schools.

Eight trios of trained young facilitators (psychologists, occupational therapists, and social workers) delivered the intervention. Facilitators received 5 days of training and weekly supervision from senior clinicians. Students received a workbook with main messages and examples. To ensure treatment integrity, a detailed operational manual was provided to facilitators, training and supervision sessions were conducted, and 10% of the sessions were evaluated by an independent ob-
server. Teachers had no involvement other than rare requests to assist with the discipline of the class. Previous studies suggested that the presence of teachers could inhibit students from sharing their experiences.17

Control
A group of students serving as controls received the standard curriculum, which included 1 hour weekly of class assembly during which problems could be discussed. If the active intervention proved effective, we offered to implement it in all control schools after completion of the trial.

Outcome Measures
The primary outcome measure was the Beck Depression Inventory–II (BDI-II),18 which provided a continuous score measured 3 months after the intervention was completed. The BDI-II is a brief self-reported depression questionnaire previously used among adolescents in Chile.19 Further criterion validation of this scale against a clinical interview was performed to establish cutoff scores according to a clinical diagnosis of depression. Although the optimal cutoff point for the whole sample was 17 or more, there were important differences between sexes, with the cutoff point for boys (≥14) being much lower than that for girls (≥20). Several other depression scales were tested in the formative phase, but BDI-II performed as well as, if not better than, other scales.

The secondary outcome measure was the Revised Child Anxiety and Depression Scale adapted from the Spence Child Anxiety Scale.20 This scale consists of 5 subscales, but we excluded the Depression and Separation Anxiety subscales because depression was measured with the BDI-II and separation anxiety was regarded as less important for this age. Additional measures included were the Personal Failure Subscale of the Children's Automatic Thoughts Scale21 and 5 subscales of the Short Form of the Social Problem-Solving Inventory–Revised22 scale. All outcomes assessed at 3 and 12 months, as well as at baseline, were determined via self-completed questionnaires. Researchers involved in the administration of these measures were frequently rotated and kept blinded to school allocation arms of the trial.

Sample Size
We aimed to find an effect size of approximately 0.4 SD as in previous successful trials.5,7 An individually randomized trial with 2 arms would require 376 individuals to detect this difference with 90% power and a 2-sided 1% α level. This size was inflated to allow for clustering and noncollection of primary outcome data. Using previously gathered local data, we calculated an intraclass correlation coefficient of 0.04 (95% CI, 0.04–0.05) for negative emotionality.23 Schools in the study had 2 to 4 participating classes with approximately 40 students per class and a mean cluster size of 80. Using a formula for infla-
tion of sample size in cluster randomized trials with unequal cluster sizes, we estimated that 2634 students from 203 schools were needed to maintain 90% power for the primary analysis.

**Statistical Analysis**

We used descriptive statistics to assess balance between the trial arms at baseline. The primary between-group analysis of covariance was carried out on an intention-to-treat basis for 3-month BDI-II scores using multivariable mixed-effects regression to account for the clustered nature of the data and to adjust for baseline BDI-II scores and randomization variables. Secondary analyses comprised BDI-II scores at 12 months and secondary outcomes at 3 and 12 months. We used repeated-measures mixed-effects regression models to investigate convergence and divergence between trial arms over time. We conducted preplanned subgroup analyses for the primary outcome using interaction terms in the regression models between randomized arm and the following baseline variables: sex, age (<14, ≥14 years), symptom severity (defined as a BDI-II score ≥20 for girls and ≥14 for boys), and contact with a psychologist or psychiatrist during the 3 months before the trial. Sensitivity analyses were conducted to assess the effects of missing data using multiple imputations. Results with and without imputed data were conducted to assess the effects of missing data using multiple imputations. Results with and without imputed data were virtually the same, and so the main results presented are from 2-level models based only on observed data. For the primary outcome only, we used instrumental variable regression models to determine whether any treatment effect was associated with the number of sessions attended, defined as 6 or more sessions, and weighted using inverse probability weights constructed using baseline BDI-II scores and randomized group.

**Results**

There were 85 schools in the sampling frame, 22 of which participated in the trial. A total of 2512 secondary school students from 66 classes in these schools participated. The Figure shows the flow of schools and students in the study. The number of students recruited in each arm was comparable. All but 2 students in each arm consented to participate in the study. Primary outcome data at 3 months were available for 82.1% (n = 2063) of participants, with completion slightly lower in the intervention (78.3%) compared with the control (85.7%) arms. Data were collected at 12 months from 77.1% (n = 1936) of participants.

The trial arms were well balanced at baseline (Table 1). There were more boys than girls, and mean age for the whole sample at baseline was 14.5 years, with 77.1% aged 14 or 15 years. The prevalence of likely clinical depressive disorder at baseline using our sex-specific cutoff points was 35% and 28% among girls and boys, respectively.

There was no evidence of any clinically important differences between the intervention and control arms in BDI-II scores at 3 months (adjusted difference in means, −0.19; 95% CI, −1.22 to 0.84; P = .72) or at 12 months (Table 2). The adjusted difference in the primary outcome at 3 months between trial arms was −0.15 (95% CI, −1.12 to 0.81; P = .38), age (0.01; −1.99 to 1.99; P = .97), gender (0.01; −2.53 to 2.53; P = .97), and baseline BDI-II score (−0.54; −2.81 to 1.70; P = .43), or previous contact with a psychologist (0.60; −5.20 to 6.40; P = .84) or a psychiatrist (−2.64; −13.22 to 7.95; P = .63).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Control Arm (n = 1289) (51%)</th>
<th>Intervention Arm (n = 1219) (49%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, No. (valid %)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>699 (54.2)</td>
<td>694 (56.9)</td>
</tr>
<tr>
<td>Female</td>
<td>590 (45.8)</td>
<td>525 (43.1)</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td>14.5 (0.9)</td>
<td>14.5 (0.9)</td>
</tr>
<tr>
<td>No. of classes per school, No. (valid %)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2 (18.2)</td>
<td>5 (45.4)</td>
</tr>
<tr>
<td>3</td>
<td>6 (54.5)</td>
<td>1 (9.1)</td>
</tr>
<tr>
<td>4</td>
<td>3 (27.3)</td>
<td>5 (45.4)</td>
</tr>
<tr>
<td>School Social Deprivation Index, mean (SD)</td>
<td>0.85 (0.1)</td>
<td>0.85 (0.1)</td>
</tr>
<tr>
<td>School attendance rate, mean (SD)</td>
<td>0.81 (0.1)</td>
<td>0.80 (0.1)</td>
</tr>
<tr>
<td>BDI-II score, mean (SD)</td>
<td>13.5 (10.4)</td>
<td>13.4 (10.1)</td>
</tr>
<tr>
<td>CATS score, mean (SD)</td>
<td>10.3 (7.9)</td>
<td>9.9 (7.9)</td>
</tr>
<tr>
<td>RCADS score, mean (SD)</td>
<td>19.8 (8.7)</td>
<td>19.7 (8.4)</td>
</tr>
<tr>
<td>Suicidal thoughts, No. (valid %)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1032 (80.1)</td>
<td>993 (81.6)</td>
</tr>
<tr>
<td>Yes, &gt;14 d before study</td>
<td>181 (14.1)</td>
<td>165 (13.6)</td>
</tr>
<tr>
<td>Yes, in past 14 d</td>
<td>75 (5.8)</td>
<td>59 (4.8)</td>
</tr>
<tr>
<td>Cigarette smoking, at 30 d, No. (valid %)</td>
<td>495 (38.5)</td>
<td>486 (39.9)</td>
</tr>
<tr>
<td>Alcohol, past 30 d, No. (valid %)</td>
<td>398 (31.0)</td>
<td>381 (31.3)</td>
</tr>
<tr>
<td>No. consuming ≥5 drinks, past 30 d, No. (valid %)</td>
<td>236 (18.4)</td>
<td>246 (20.2)</td>
</tr>
<tr>
<td>No. of occasions consuming ≥5 drinks, past 30 d, No. (valid %)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2 Times</td>
<td>153 (11.9)</td>
<td>160 (13.3)</td>
</tr>
<tr>
<td>≥3 Times</td>
<td>83 (6.5)</td>
<td>86 (7.1)</td>
</tr>
<tr>
<td>Cannabis use, past 30 d, No. (valid %)</td>
<td>163 (12.7)</td>
<td>182 (15.0)</td>
</tr>
<tr>
<td>Lifetime GP contact for mental health, No. (valid %)</td>
<td>197 (15.3)</td>
<td>187 (15.4)</td>
</tr>
<tr>
<td>Missing school for mental health reasons, past 3 mo, No. (valid %)</td>
<td>64 (11.0)</td>
<td>49 (10.3)</td>
</tr>
<tr>
<td>Psychiatrist contact, past 3 mo, No. (valid %)</td>
<td>19 (1.5)</td>
<td>21 (1.7)</td>
</tr>
<tr>
<td>Psychologist contact, past 3 mo, No. (valid %)</td>
<td>53 (4.1)</td>
<td>48 (3.9)</td>
</tr>
</tbody>
</table>

Abbreviations: BDI-II, Beck Depression Inventory–II; CATS, Children’s Automatic Thoughts Scale; GP, general practitioner; RCADS, Revised Child Anxiety and Depression Scale.
There was no evidence that any difference between the intervention and control arms in BDI-II score varied over time, with an interaction coefficient of −0.14 (95% CI, −1.20 to 0.93; P = .80).

The mean (SD) number of sessions attended was 8.4 (3.4), with 80.5% of students attending at least 6 sessions. The instrumental variable analysis of the primary outcome showed no evidence of any treatment effect, with the estimated difference between intervention and control being −0.07 (95% CI, −1.22 to 0.84) with adherences defined as students having received 6 or more sessions. Assessing the additional intervention effect per sessions attended gave an estimated difference of −0.01 (95% CI, −0.13 to 0.12; P = .91).

Discussion

To our knowledge, this is the first large randomized clinical trial from Latin America of a universal school-based intervention aimed at reducing depressive symptoms. We found no evidence of this intervention being better than usual care in any of the main outcomes, with 95% CIs ruling out meaningful clinical effects. The study addressed many of the shortcomings of previous trials of depression prevention interventions throughout the world. The population involved showed high levels of depressive symptoms, with almost one-third likely clinical cases of depression and a sizeable proportion expressing suicidal ideas. The intervention was carefully developed, based on the best evidence available, and had good levels of attendance. The study used robust methodology, was a meticulously designed intervention with good adherence, had low attrition rates for the primary analysis, and had a reasonably long follow-up period.

Among the limitations, no psychiatric interviews were conducted because it was impractical given the size of the study. Moreover, we aimed to implement a pragmatic trial, and if this intervention were to be extended to a larger study population, only questionnaires would be used. Although the BDI-II may have been designed for use primarily with clinical populations, there is evidence that it is sensitive to change with general adolescent populations in Latin America and elsewhere. Delivering booster sessions proved difficult, since a large proportion of the students changed schools. The population under study was selected in terms of their low socioeconomic background because this was our population of interest. We did not include a placebo arm because of the challenges in designing an appropriate comparison, the increased complexity and costs of a 3-arm study, and the pragmatic focus of the trial. It was not possible to have more intensive fidelity checks given the extent of this trial as well as other practicalities. Finally, although there was some differential attrition across arms in the final follow-up assessment, results remained unaltered when imputing missing values.

Schools are an attractive setting to introduce preventive or early treatment interventions to help young people cope better with life’s vicissitudes. Unsurprisingly, school mental health programs are being delivered throughout the world, but most of them have never been evaluated adequately. Among those universal, school-based psychological interventions to pre-
vent or reduce depressive symptoms that have been evaluated, effects have been absent or small regardless of the therapeutic approach or modality to deliver the intervention. Better results have been found with targeted interventions and when treating clinical populations, suggesting that psychological interventions under certain conditions may be effective with adolescents. So why do school-based psychological interventions to prevent or reduce depressive symptoms not seem to work when delivered to whole classrooms?

First, classroom-based interventions may not reach the intensity needed to achieve effects on depression outcomes. It might not be possible to teach complex skills in large groups and during limited time. Most previous universal school-based trials were conducted in groups of 15 or fewer students. A universal approach was chosen because students with subthreshold symptoms might benefit and to avoid potentially stigmatizing those who would have been targeted for an intervention. Second, achieving improvements when most adolescents initially have few symptoms is challenging. However, the unexpected high symptom levels at baseline in this population somehow undermine this argument. In addition, we found no evidence of interactions between trial arms and baseline severity of symptoms, but admittedly, the statistical power for subgroup analyses in trials is generally low. Third, the intervention may have been of poor quality or unappealing to recipients. A long and rigorous process was undertaken to develop an appealing intervention for this audience. We tested each component, consulted widely with experts, and carried out a large pilot study to reduce this possibility. Students were satisfied with the intervention (data not shown); this explains the good attendance rate. Fourth, facilitators with limited experience and training may have reduced the potential effectiveness of the intervention. However, our facilitators were carefully selected, trained, and regularly supervised by experienced professionals. Variability in therapists’ effects is a well-documented phenomenon in psychotherapy. However, this variability was difficult to estimate because facilitators delivering the intervention in each class differed for practical reasons. Nonetheless, we found that the results were fairly consistent across schools and classes. Fifth, it is difficult to deliver school interventions without some degree of flexibility to accommodate for unanticipated events. For instance, we were unable to deliver the booster sessions as planned because a large proportion of the students changed schools or classes and because a student strike lasting several weeks occurred during implementation of the regimen. Finally, school-based interventions targeting specific problems in populations with marked social deprivation may prove insufficient to change mental health or other social outcomes. More comprehensive interventions may be needed that include access to specialized health care services for those with severe symptoms. However, this option requires highly specialized resources—something difficult to accomplish in low- to middle-income countries.

Recent reviews ascertained that most universal school-based studies to prevent depressive symptoms are of low or moderate methodologic quality. In the most recent review only 4 studies showed some, albeit short-term, effects, but all were limited by significant methodologic problems. There was also notable heterogeneity regarding content and delivery of the intervention, group size, and methods of analyses, making it difficult to draw any useful conclusions from these more successful studies. Among recent depression prevention studies not included in the review, all 4 reported no clinically meaningful effects. Two of these recent studies were small but undertaken in low- to middle-income countries and the other 2 studies conducted in high-income countries were the largest of this kind ever undertaken. The samples of these studies added to our study would result in a sample larger than that from all previous studies included in the most recent systematic review. The null results from these methodologically robust and large studies from a range of settings are likely to eliminate the small but positive results found in the recent meta-analysis.

In conclusion, given the lack of effect in this and other universal school-based trials for depressive symptoms, it is legitimate to ask whether we should continue to invest in these programs. This is important because there are many mental health school-based programs under way throughout the world, including in low- to middle-income countries, with no evidence to support their effectiveness. While we continue to search for evidence-based programs, a note of caution must precede the introduction or continuation of mental health school-based interventions.