JOURNAL OF CHILD AND ADOLESCENT PSYCHOPHARMACOLOGY
Volume 22, Number 1, 2012

Pp. 80–90
DOI: 10.1089/cap.2011.0074

Treatment of Youth Depression in Primary Care Under Usual Practice Conditions: Observational Findings from Youth Partners in Care

Kenneth B. Wells, M.D., M.P.H.,[1] Lingqi Tang, Ph.D.,[1] Gabrielle A. Carlson, M.D.,[2] and Joan Rosenbaum Asarnow, Ph.D.[1]

Abstract

Objectives: The effectiveness of treatments for youth depression in primary care, under usual practice conditions, is largely unstudied. This study aims at estimating the effect of ''appropriate treatment,'' defined as treatment that approximates guideline standards, on clinical outcomes for depressed primary care youth patients by using observational analyses from a randomized trial.

Methods: Participants were 344 youths aged 13–21 enrolled in the Youth Partners in Care trial. Youths screening positive for depression from six primary care practices in five different health care organizations were randomly assigned to either (1) usual care enhanced by provider education on depression evaluation and management, or (2) a quality improvement (QI) intervention designed to improve access to antidepressant medications and/or cognitive behavior therapy for depression; usual practice conditions otherwise applied. Observational analysis was conducted on the effects of appropriate treatment (antidepressant medication use by algorithms or 6 or more psychotherapy visits) on severe depression (Center for Epidemiologic Studies-Depression score ≥ 24) at 6 months. Selection into treatment is accounted for by using instrumental variables analysis, with randomized QI intervention status as the instrument.

Results: At 6 months, youths receiving ''appropriate treatment,'' compared with others, were significantly less likely to have severe depression (10.9% vs. 45.2%, $p < 0.0001$). Similar findings were observed among youths with depressive disorders and sub-syndromal depressive symptoms, and among Latino and other youths.

Conclusions: Among depressed primary care youths, care that approximates guideline standards but retains leniency substantially reduces the likelihood of severe depression at 6 months. Such findings apply to youths with or without depressive disorder, and among Latino youth.

Introduction

Depressive disorders and symptoms are common in youth, with disorders affecting 6% and symptoms 28% in a year; and depression is associated with impairment in functioning, increased risk of suicide, and adverse outcomes such as substance abuse (Kessler 2002; Birmaher et al. 2007; Eaton et al. 2009). Many depressed youths remain untreated, especially minority youths, and there is evidence of poor quality and outcomes of care in community-based settings, relative to expectations from clinical trials (Horwitz et al. 1992; Jensen et al. 1999; U.S. Department of Health and Human Services 1999; Weisz et al. 2005).

Evidence-based treatments supported by clinical trials include cognitive-behavior therapy, interpersonal psychotherapy, and certain selective serotonin reuptake inhibitors (SSRIs) (Emslie et al. 1997; Clarke et al. 1999; March et al. 2004; Mufson et al. 2004; Vitiello and Swedo 2004; Wagner et al. 2004a, 2004b; Brent et al. 2009). Although there is conflicting evidence (Goodyer et al. 2008), one large study suggests that combined cognitive behavior therapy (CBT) and medication or medication alone may be superior to CBT for the treatment of moderate to severe adolescent major depression (March et al. 2004), and that adolescents resistant to an initial SSRI improve significantly more with a switch to another SSRI or venlafaxine plus CBT, when compared with a medication switch alone

[1]Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, California.
[2]Department of Psychiatry, Stony Brook University, Stony Brook, New York.

The analyses for this article were funded by P30 MH082760. The original study was funded by grant HS09908 from the Agency for Healthcare Research and Quality, with additional support from NIMH grant P30MH068639; from the University of California, Los Angeles-Jane and Terry Semel Institute and David Geffen School of Medicine (Drs. Wells, Tang, and Asarnow); and UCLA School of Public Health and RAND Corporation, Santa Monica CA (Dr. Wells); and Stony Brook University School of Medicine (Dr. Carlson).

(Brent et al. 2009). Due to concerns about safety and efficacy of SSRIs in youth (Vitiello and Swedo 2004), antidepressants have a black box warning (U.S. Food and Drug Administration 2004). Poor performance of CBT alone compared with medication or combined treatment in some studies (March et al. 2004), and poor benefit to risk ratios for medications other than fluoxetine, given findings of unpublished pharmaceutical trials (Vitiello and Swedo 2004), has led to uncertainty about treatment effectiveness.

In this era of health-care reform, a major challenge in the field involves identifying strategies for improving patient outcomes under routine practice conditions. Despite advances in the efficacy literature, there are relatively few demonstrations in the literature of improved patient outcomes when treatments are delivered under routine naturalistic practice conditions. Estimates of effectiveness of treatments as delivered under naturalistic practice conditions have policy relevance, because they clarify the outcomes achieved in the community (Sturm and Wells 1995; Sturm 2006). For example, reimbursement mechanisms might not support a full CBT course, and patients or families may not use medications as recommended. Such conditions could lead to worse outcomes compared with efficacy studies if treatment adherence is poor, or better outcomes if treatments are more effective when they match patient preference.

Estimates of the effectiveness of treatments delivered under usual practice conditions are difficult to achieve. Most experimental designs alter many features of usual care (UC), whereas estimates from observational studies preserve UC conditions but are subject to strong selection biases due to unmeasured sickness differences (Sturm 2006). One approach to address such selection bias in observational analysis is instrumental variables (IV) analysis, an econometric method that relies on identifying an external factor which affects the probability of treatment but does not directly affect outcome (Sturm 2006). This method has been used in observational effectiveness studies of medical/surgical treatments (McLellan et al. 1994), and to demonstrate that appropriate treatment for depressed adults improves clinical and employment outcomes (Schoenbaum et al. 2002). The instrument in the Schoenbaum et al. (2002) study was randomized quality improvement (QI) intervention status, a strategy that helps assure that the assumptions underlying the method are met (Heckman 1996; Schoenbaum et al. 2002).

In this study, we apply a similar strategy to data from Youth Partners in Care (YPIC) (Asarnow et al. 2005) to estimate the effect of probable appropriate treatment on outcomes for depressed youth patients receiving primary care. YPIC was designed to determine the effectiveness of a QI intervention for depressed youth in primary care. The QI intervention increased access to evidence-based depression treatment (primarily cognitive-behavior therapy for depression and/or antidepressant medication) by using ''a randomized encouragement design'' where resources were available to the providers, but the choice regarding whether to use the treatment resources was made by providers and patients/families. Compared with UC enhanced by provider education on depression evaluation and treatment (primarily pharmacotherapy, the main treatment option within usual primary care), the YPIC QI intervention improved 6-month depression and quality of life/functioning outcomes, increased use of any specialty mental health visit, and of treatment, particularly psychotherapy/counseling, but did not significantly affect use of medication (note that this study was conducted before the Black Box warning for antidepressant medications). Early improvements associated with the QI intervention at 6 months appeared to shift youths towards healthier pathways through 12 and 18 months of follow-up (Asarnow et al. 2009). Unlike a traditional clinical trial, under the YPIC QI intervention, patients and providers were encouraged but not required to consider evidence-based treatments and given trainings and materials to support their use. In both study conditions, primary care providers received education in depression care but applied any treatments under usual practice conditions and constraints; in particular, the study did not reimburse treatments and services. Since usual practice conditions were largely maintained, the randomized intervention status is a good candidate as an instrument in IV analysis to examine the impact of use of appropriate treatment on outcomes in a secondary observational analysis using data from the randomized trial of QI. This study, thus, expands on previous results on the effectiveness of the QI intervention (i.e., having the full QI resources vs. UC enhanced by provider education) to address the question: How effective is depression treatment meeting at least minimal criteria for appropriate care, when it is provided under usual practice conditions; that is, when patients and providers actually choose treatments and usual conditions apply?

In addition, this study included youths who had depressive disorder, who might benefit from treatment, and those with depressive symptoms but without disorder, who might benefit from prevention or early intervention (Clarke et al. 1995). Treatment guidelines focus on treatment for disorder (Birmaher et al. 2007); so, an exploration of effectiveness of treatment among those with symptoms only is an important complementary opportunity. Since YPIC had a large sample of Latino youths, the study also offers an opportunity to explore effectiveness of care for depression in this under-studied population. The adult PIC study found that appropriate treatment was effective for depressed minorities (Miranda et al. 2004). We hypothesized that those who received appropriate treatment, defined as having at least six psychotherapy/counseling sessions or use of antidepressant medication in 6 months, would be less likely to have severe depression at 6-month follow-up, compared with those without such care, using IV analysis. We hypothesized that treatment-outcome relationships would not be significant by using traditional analyses. Based on findings for adults, we thought that conclusions would be similar for depressed youths with and without depressive disorder, and among Latino and other youths (Clarke et al. 1995; Wells et al. 2000).

Methods

YPIC is a multisite randomized, controlled trial comparing the effectiveness of a QI intervention for depressed youth, with enhanced UC (provider education only), in primary care (Asarnow et al. 2005). The study protocol was approved by the institutional review boards of participating organizations. Youth under the age of 18 provided written assent, and parents or legal guardians provided written informed consent, whereas youth 18 or older provided written consent. The design and methods of the study are described elsewhere (Asarnow et al. 2005, 2009).

Organizations and Clinicians

Six practices from five organizations were purposively selected to include public sector, managed care, and academic primary-care practices. Fifty-two primary care providers enrolled at participating clinics.

Patients

Patients of enrolled providers were recruited through screening a consecutive sample of clinic visitors to enrolled providers, aged

13–21. Youth were eligible if they met either of the 2 criteria: (1) endorsed one or more ''stem items'' for major depressive or dysthymic disorder from the 12-month Composite International Diagnostic Interview (CIDI-12, Core Versions 2.10) (World Health Organization 1997) as modified for diagnostic criteria for adolescents; plus had a Center for Epidemiological Studies-Depression Scale (CES-D) of 16 or greater; or (2) a CES-D score of 24 or greater (Radloff 1977). Exclusion criteria included those not speaking English and those with an enrolled sibling. The recruitment period was 1999–2002.

Of 4,750 youth eligible for screening, 4,002 completed screening (84%), of whom 1,034 (26%) met eligibility criteria; of these, 418 (40%) enrolled, completed baseline assessment, and were randomized. Reasons for not enrolling include problems contacting (259), refused or failed to complete consent (289), or baseline (68). After baseline, patients were randomly assigned to the QI intervention or enhanced UC, using a computerized random number generator. Randomization was stratified by site. Participants recruited from the same clinician (409) were randomized in pairs based on enrollment date, whereas 9 sole participants of an enrolled clinician were randomly assigned into one of two conditions. Among 418 youth enrolled, 344 (82%) completed 6-month follow-up, with no significant difference in completion by intervention status.

Study Randomized Intervention Conditions (IV)

Enhanced UC

UC was enhanced by providing all primary care clinicians with training and educational materials on depression evaluation and treatment and implementing treatment plans while considering available options and patient/family preferences/cultural background. The training reviewed medication management based on the Texas Medication Algorithms for Major Depressive Disorder, emphasizing certain SSRIs as the first choice (Asarnow et al. 1996; Hughes et al. 1999). Providers had usual access to treatments for all patients, but study-trained care managers were available only to patients randomized to QI.

QI condition

The QI intervention was modeled on the adult Partners in Care study (Wells et al. 2000), modified for youth. Components included (1) expert practice leaders at each site who adapted and implemented the intervention; and (2) care managers at the sites who supported clinicians with patient evaluation, education, treatment initiation and follow-up, services linkages, and provided manualized CBT for depression (Clarke et al. 1990; Asarnow et al. 1999). Care managers had a masters or doctoral degree in a mental health or nursing field. They were available to follow patients for 6 months to coordinate care, assist with treatment, and provide CBT as needed. In addition to the primary care clinician training in depression evaluation and management offered across conditions, the study provided training, written manuals, and consultation to support model fidelity and case consultation on the care manager and cognitive-behavior therapy components of the intervention. The study paid for care managers' time, which was available to patients with QI and parents/guardians without co-pay. Sites were requested to waive co-pay for the first care manager visit but generally opted to do so for most or all such visits. The study CBT was based on Adolescent Coping with Depression Course, for individuals and groups (Clarke et al. 1990; Asarnow et al. 1999). This CBT includes an overview session, three 4-session modules on activities and social skills, cognition, and communication and problem-solving, and a final session on relapse prevention.

Measures

At baseline, patients provided information on age, sex, and ethnicity/cultural background. At baseline and 6-month follow-up, youth completed the CIDI Diagnostic Interview Schedule (World Health Organization 1997) and Mental Health Inventory 5 (MHI-5) (Berwick et al. 1991) and reported services use during the previous 6 months by using the adapted Service Assessment for Children and Adolescents (Horwitz et al. 2001; Asarnow et al. 2005). CES-D was collected at 6-month follow-up (Radloff 1977). The outcome measure is an indicator for severe depression, defined as CES-D ≥ 24. Other data collected by the study are not included in these analyses.

Probable appropriate treatment

In quality of care research, indicators are constructed of appropriate treatment that roughly follow guideline standards but are broadened to give clinicians and patients benefit of the doubt. For example, a patient may start treatment before a given study assessment. Sturm and Wells (1995) used decision analysis to demonstrate that among depressed adults, those receiving at least four counseling visits or 2 months of antidepressant medication had better outcomes. In PIC, a similar indicator significantly improved health status and employment at 6 months, using IV analysis (Schoenbaum et al. 2002).

We formulated an indicator of probable appropriate treatment for depressed youths: (1) use of any SSRI or venlafaxine or buproprion (following guideline recommendations) (Hughes et al. 1999); or (2) having 6 or more specialty counseling visits, during 6 months of follow-up. We corrected spelling errors in medication names. We did not consider trazodone an antidepressant, as dosage data suggested that it was used for sleep. For sensitivity analyses, we created alternatives that required daily medication use for 2 or more months and applied minimum dosage criteria; specified the minimum number of counseling visits as 4 or 12; and further broadened the definition of appropriate care to include use of any antidepressant medication, counseling visit, or mental health specialty visit as a measure of access to treatment or specialty evaluation, with no change in conclusions. These various indicators can be conceptualized as varying in terms of the breadth versus strictness of the definition of care/treatment, with the ''appropriate treatment indicator'' representing the stricter/lower bound definition of treatment/care and the ''access to treatment or evaluation'' indicator representing a broader/upper bound definition. Given controversies over benefits and risks of medications for depressed youth, we also specified an indicator using data on counseling and ignoring medication use (see Table A1 in Appendix). This was for an exploratory sensitivity analysis under the hypothetical assumption that if medications are not effective, then information on their use can be ignored in defining overall appropriate care.

Data analyses

We used univariate and bivariate analyses to describe the sample and compare patients with and without probable appropriate treatment over the 6 months of follow-up.

We used IV analysis (Angrist et al. 1996; Heckman 1996; Imbens and Rubin 1997; Imbens and Rosenbaum 2005; Sturm 2006;

Rosenbaum 2010) to account for selection effects. IV analysis relies on identifying an instrument that predicts the probability of treatment, but has no independent effect on outcomes. We used randomized intervention status as the instrument, which meets the conditions for valid instruments: (1) random assignment (because QI vs. UC is randomized); (2) the instrument affects treatment use (''nonzero average causal effects'') (as shown in Asarnow et al. 2005); (3) the instrument is unlikely to affect outcomes except through treatment (''exclusion restriction''), because the intervention provides resources to encourage youth depression treatment and was randomly assigned at the provider level, and the treatment indicator is broadly defined to include the major care components encouraged by the intervention; (4) outcomes for one patient are unlikely to be affected by outcomes for another (''stable unit treatment value''), as each patient should respond to treatment individually, and outcomes for individuals are independently assessed; and (5) patients improving in UC would also improve under the intervention (''monotonicity''), because similar treatments are potentially available, and some encouragement is provided through provider training even in UC (Angrist et al. 1996). Following recommendations in the literature (Bhattacharya et al. 2006), we conducted IV analysis by simultaneous estimation using bivariate probit regression to jointly model probable appropriate treatment and severe depression outcome, explicitly taking into account the correlation between the unobserved factors in two equations (see Table A2 in Appendix). As a sensitivity analysis, we also conducted IV analysis by using two-step probit regression models with no change in conclusions (see Table A3 in Appendix). In the bivariate probit model, we included the following independent variables in predicting ''probable appropriate treatment'' versus ''no treatment'': The IV (intervention status), age, gender, and baseline MHI-5 score. In the equation for depression outcome, we included the same covariates but replaced the IV with the estimated probable appropriate treatment. For sensitivity analyses, we also included ethnic status (Latino, African American, non Hispanic white, and other); and use of another language than English at home, with no change in main conclusions. To estimate the bivariate probit model, we used QLIM procedure in SAS/ETS V9.2. As a sensitivity analysis, we also fit 'Seemingly unrelated bivariate probit model' by using biprobit command in STATA version 11, with no change in conclusions. Since the correlation between the error terms of the two probit regressions (rho) was close to 1, we conducted sensitivity analyses for values of rho from .80 to .99 with no change in conclusions. To be conservative, we use the BOUNDS statement in PROC QLIM to impose boundary constraint to rho ≤ 0.85, the value of the correlation in PIC (0.85). We conducted one set of analyses for the full sample, as the main analysis, and sets of stratified analyses, considered exploratory, for youth with depressive symptoms with and without current depressive disorder according to CIDI criteria, and for Latino and other (mixed ethnicity) youth. Our inference statistics are based on the significance of the estimated coefficients of the bivariate probit model. We used the parameter estimates from the regressions and each individual's actual value for covariates to generate predicted values under the scenario that the patient received probable appropriate treatment, and then under the scenario without probable appropriate treatment. After that, we averaged the predictions across individuals under each scenario (Korn and Graubard 1999).

Although patients were clustered within providers, intraclass correlation coefficients are close to zero (<0.01); so, we report results unadjusted for cluster effects. In sensitivity analyses, we used non-response weighting to account for missing data for the 18% of patients who did not complete 6-month follow-up (Asarnow et al. 2005). Weighted and unweighted bivariate probit models did not differ in conclusions, so we report unweighted analyses.

Results

Ethnic/racial background of patients differed significantly for those with appropriate treatment versus those without probable appropriate treatment. As shown in Table 1, the pattern of results suggests that minorities/non-English speakers were underrepresented among those receiving appropriate treatment. In addition, consistent with a selection bias where sicker youths tend to receive more treatment, youths receiving appropriate treatment had significantly worse baseline mental health status as indicated by significantly higher rates of depressive disorders and higher CES-D scores and rates of severe depression (CES-D ≥ 24) and lower MHI-5 scores.

Of youth providing follow-up data, 34 reported using any eligible antidepressant medication (any SSRI, venlafaxine, or buproprion). Twenty-five used any SSRI, 6 venlafaxine, and 5 buproprion (of which one also used an SSRI and one venlafaxine). Fifty-eight youths had six or more counseling visits. Of these, the median number of sessions in 6 months was 15; whereas among those with fewer than 6 visits, the median was 0. Among 34 youth who used an eligible antidepressant medication, 24 (71%) had six or more counseling visits. Of all youths receiving an appropriate antidepressant, only two (8%) used dosages below the recommend minimum threshold, and only one used that medication for less than 2 months.

For the whole sample, the results of the bivariate probit IV analysis confirm a statistically significant effect of the ''instrument'' (intervention condition status) on the probability of appropriate treatment (probit coefficient = 0.32 [standard error = 0.13]; $t = 2.46$ [$p = 0.01$]) (see Table A2 in Appendix). Table 2 shows the main findings of the estimated effect of probable appropriate treatment on severe depression, by IV (bivariate probit) analysis, for the overall sample and stratified by initial disorder status and by ethnic grouping. For each analysis, the estimated effect of appropriate treatment is highly statistically significant (each $p < 0.0001$), thus indicating that appropriate treatment significantly improved outcome overall, and among those with and without current disorder at baseline (given depressive symptoms), and among Latinos and non-Latino youth. Further, the estimated average effect sizes are large. For example, the predicted percentages with severe depression for the overall sample are 45.2 (confidence interval [CI], 39.8–50.6) for those without appropriate treatment falling to 10.9 (CI, 6.8–14.9) for those with appropriate treatment. The difference in mean estimated percent severely depressed is about 30 percentage points overall within each subgroup, but the absolute levels are higher for those with depressive disorder (i.e., 55.9 vs. 20.5%) compared with those with symptoms but no disorder (37.0 vs. 3.9%). In exploratory analyses, we found that when limiting the definition of appropriate treatment to 6 or more counseling visits (ignoring information on medication use), the effect of appropriate treatment on outcomes for those with symptoms only remained significant ($t = -6.48$, $p < 0.001$) but was insignificant among those with disorder ($t = -1.15$, $p = 0.25$) (see Table A1 in Appendix). This suggests that for primary care youth with depressive disorders, information on use of medication was necessary to define an overall indicator of appropriate care that improved 6-month outcome. (See Table A2 in Appendix for results for the full model for the main IV analysis including controlled covariate effects.)

Table 1. Patient Characteristics for Analytic Sample and by Appropriate Treatment Status over 6-Month Follow-Up[a]

Characteristics	*Overall (n = 342)*	*Not appropriate treatment (n = 274)*	*Appropriate treatment (n = 68)*	*Statistic*[b]	p-*Value*[b]
Demographics					
Age, mean (SD), years	17.0 (2.1)	17.1 (2.1)	16.9 (2.1)	$t_{[340]} = -0.58$	0.57
Female sex, No. (%)	265 (77.5)	209 (76.3)	56 (82.4)	$Chisq_{[1]} = 1.15$	0.28
Race, No. (%)				$Chisq_{[3]} = 14.96$	< 0.01
White	40 (11.7)	23 (8.4)	17 (25.0)		
African American	59 (17.3)	48 (17.5)	11 (16.2)		
Hispanic/Latino	224 (65.5)	188 (68.6)	36 (52.9)		
Other	19 (5.6)	15 (5.5)	4 (5.9)		
Language other than English spoken at home, No. (%)	231 (67.5)	195 (71.2)	36 (52.9)	$Chisq_{[1]} = 8.26$	< 0.001
Baseline health status					
MHI-5 score, mean (SD)	19.4 ± 4.9	19.8 ± 4.8	17.6 ± 5.0	$t_{[340]} = -3.36$	< 0.001
CIDI diagnosis of Major Affective Disorder, No. (%)[c]	144 (42.1)	98 (35.8)	46 (67.6)	$Chisq_{[1]} = 22.71$	< 0.001
Outcomes at 6 months					
CES-D, mean (SD)[b]	19.8 (12.5)	19.1 (12.1)	22.6 (13.9)	$t_{[340]} = 2.07$	0.04
CES-D ≥ 24, No. (%)	121 (35.38)	90 (32.85)	31 (45.59)	$Chisq_{[1]} = 3.87$	0.05
Number of psychotherapy or counseling visit, mean (SD), median	4.3 (10.6), 0.0	0.2 (0.8), 0.0	21.0 (15.1), 22.0	$t_{[339]} = 22.78$	< 0.001
Psychotherapy or counseling visits ≥ 6, No. (%)	58 (17.0)	0	58 (86.6)		
Use antidepressant medication, No. (%)	34 (9.9)	0	34 (50.0)		

Appropriate treatment: at least six psychotherapy or counseling sessions or use of antidepressant medication in 6 months. Range of possible scores: 0–60 (CES-D), 5–30 (MHI-5).

[a]Baseline enrollment $n = 418$. Attrition at 6 months resulted in a 6-month sample size of 344. Due to missing data, sample size = 343 for ''any antidepressant medication,'' $n = 341$ for ''psychotherapy or counseling visit,'' and $n = 342$ for ''appropriate treatment.''

[b]Test of difference between appropriate treatment and no appropriate treatment groups.

[c]Major depression, dysthymic disorder, or bipolar in past year.

MHI-5 = Mental Health Inventory; CIDI = Composite International Diagnostic Interview; CES-D = Center for Epidemiological Studies-Depression Scale.

Consistent with a selection effect, when traditional probit regression analysis with covariates was used instead of IV analyses, the effects of appropriate treatment on outcome were not statistically significant in the overall sample or any of the subsamples (e.g., for the overall sample, $t = 0.90$, $p = 0.37$), thus underscoring the value of IV analyses for estimating treatment effects when treatment is delivered under natural practice conditions. (see Table A3 in Appendix).

Using a broader indicator of appropriate treatment (any access to counseling, antidepressant medication or mental health specialty visit) generates comparable conclusions, including a significant effect of intervention status (instrument) on this treatment indicator and a highly significant effect of this treatment indicator on outcome using bivariate probit, comparable findings with the two-stage probit model, and insignificant results with traditional probit analyses (see Table A4–A6 in Appendix).

Discussion and Conclusion

Our findings suggest that primary care depressed youths aged 13–21 who receive treatment that meets at least minimum criteria for appropriate care for depression given current guidelines,

Table 2. Predicted Percent of Severe Depression at 6 Months, by Appropriate Treatment Status[a]

	Predicted % severe depression				
	Not appropriate treatment	*Appropriate treatment*	*Bivariate probit*		
	% (95% CI)	*% (95% CI)*	*Coeff*	*SE*	p-*Value*
a. Full sample ($n = 342$)	45.2 (39.8 to 50.6)	10.9 (6.8 to 14.9)	−1.23	0.13	< 0.001
b. Disorder subsample ($n = 144$)	55.9 (47.4 to 64.4)	20.5 (12.6 to 28.5)	−1.05	0.17	< 0.001
c. Symptoms only subsample ($n = 198$)	37.0 (30.2 to 43.9)	3.9 (0.5 to 7.2)	−1.54	0.22	< 0.001
d. Latino subsample ($n = 224$)	42.6 (36.0 to 49.2)	8.9 (4.2 to 13.6)	−1.24	0.17	< 0.001
e. non-Latino subsample ($n = 118$)	49.4 (39.9 to 58.8)	15.2 (7.8 to 22.6)	−1.27	0.21	< 0.001

Appropriate treatment: at least six psychotherapy or counseling sessions or use of antidepressant medication in 6 months; severe depression: CES-D ≥ 24.

[a]Results were based on treatment coefficient from bivariate probit model, adjusted for age gender and baseline MHI-5 score. In the quality improvement condition, 39 out of 170 youths (22.9%) received ''appropriate treatment'' versus 29 out of 172 youths (16.9%) in usual care.

CI = confidence interval; SE = standard error.

compared with those who do not, have a significantly reduced likelihood of severe depression, falling from 45.2% for those without appropriate treatment, to 10.9% with appropriate treatment broadly. Results were consistent using our primary indicator of ''appropriate treatment'' (at least 6 specialty counseling visits or antidepressant medication in 6 months) and with alternative definitions (daily medication use for 2 or more months with application of minimum dosage criteria; the minimum number of psychotherapy visits specified as 4 or 12; a broad ''upper bound'' indicator of treatment defined as any counseling, antidepressant medication, or mental health specialty visit). What is unique about the context for these findings is that patients and providers made their own decisions about treatments and other than paying practices for care managers' time in the QI condition, the study did not pay for services, thereby preserving aspects of ''usual'' practice conditions for treatment provision. These results apply to all patients across intervention conditions, as this is a secondary, observational analysis of outcomes of exposure to appropriate care using data from all subjects in the randomized QI trial. The IV analyses offer a form of ''as-treated'' analysis in which across intervention conditions patients receiving broadly-defined ''appropriate'' depression treatment are compared with those who have not received ''appropriate'' treatment, with adjustment for potential unmeasured selection effects that can lead to sicker patients having more treatments but worse outcomes.

In this study, most youths receiving appropriate treatment had more than the minimum care required to meet criteria for the indicator. Most youths on antidepressants used them daily for 6 months, and the median number of counseling sessions among those with 6 or more was 15, which is similar to the number of sessions for a course of evidence-based psychotherapy such as CBT. Further, most youths using antidepressants also received counseling (71%), thus indicating combined treatment. The treatment-outcome relationships we estimated were due to this full range of treatments received among those meeting at least minimum criteria for appropriate care. Analyses using data only on counseling to define appropriate care and ignoring information on use/nonuse of medications yielded non-significant effects for youths with depressive disorders. This exploratory finding underscores the importance of including medication to define appropriate treatment for more severely depressed youths and indirectly suggests that medication use is a part of what makes ''appropriate care'' effective for this group. We used lower standards for acceptable treatment than are typical for defining guideline-concordant care, because we did not observe care in adjacent periods and did not have full histories of care, and in this circumstance, it is common to give the provider the benefit of the doubt in such effectiveness analyses. With an IV analysis approach, we could only specify as many treatment effects as there were instruments (i.e., one), so we could not separate the effects of medications and therapy, specify interactions, or determine how much care is necessary to improve outcomes. Indeed, the very broad indicator of access to any treatment or specialty consultation revealed similar results, thus suggesting that research is needed on how much of the value in effective care for primary care depressed youth, where treatment rates are low, can be attributed to use of evidence-based treatments, access to any treatment, or access to specialty consultation. This study can raise but not resolve this issue.

These findings are an important complement to the main experimental findings from YPIC (Asarnow et al. 2005), which showed that a QI intervention had a modest effect on increasing both treatment rates (particularly psychotherapy/counseling) and clinical outcomes. The new observational IV analyses showing that receiving ''appropriate treatment'' substantially improves outcomes reinforce the importance of implementing interventions in primary care to increase rates of appropriate care as well as of further improving the effectiveness of QI interventions in doing so.

We note that the relatively modest improvements in treatment rates and clinical outcomes previously reported in the YPIC main experimental analyses (Campo and Bridge 2009) are typical of QI interventions that encourage but do not assign treatments in a policy environment which does not strongly encourage the use of such programs. However, such programs are likely to be of increased importance in an era of implementation of parity and health reform legislation focusing on greater integration and quality of care (Barry and Huskamp 2011). Our findings suggest that it did not take the rigorous implementation conditions of clinical trials for clinicians to provide and patients to receive treatments that substantially improved their outcomes including Latino youths, an important group for expanded coverage under health reform legislation.

The sample included patients with depressive disorder and patients with symptoms but without past-year disorder. We found a strong effect of appropriate treatment on outcome within each group. Thus, our findings may reflect a combination of treatment response among those with current disorder, early intervention among those with a previous disorder, and prevention of disorder among those with symptoms but no disorder history. For example, there is evidence that CBT is effective in preventing depressive disorder in high-risk youths (Clarke et al. 1995; Garber et al. 2009). Development of treatment approaches that encompass these diverse goals through an integrated practice strategy is an important issue for further research.

This study has important limitations. We had moderate response rates at enrollment, which somewhat limits generalizability—our findings may be more applicable to youths and families willing to address depression. Our study was limited to particular sites, and we did not have the sample to estimate site differences.

In YPIC, clinicians in both intervention arms had education in evidence-based practice. Although this feature likely shifted rates of appropriate care upward somewhat overall across both intervention arms relative to similar practices, we do not think that it shifted the effectiveness of appropriate care when provided, thus making the estimates provided here of potentially broad interest. This is especially the case because the analysis used data on depressed youth in both conditions, many youths did not receive appropriate treatment under each condition, the appropriate treatment indicator was applied across conditions, there was no CBT training in UC, and the medication criterion would be applicable across any practice setting.

Other limitations of the IV analysis approach arise if assumptions underlying the method may plausibly not hold, such as the instrument not being a strong predictor of treatment, thus resulting in under-estimation of CI for estimated treatment-outcome relationships; or when the instrument is not plausibly random with regard to outcome or the treatment indicator does not capture the instrument's effects on outcomes, thus leading to bias (Imbens and Rubin 1997; Imbens and Rubin 2005; Rosenbaum 2010). The randomization of the intervention condition in YPIC supports its independence with regard to influencing outcome except through treatment effects. We included sensitivity analyses with a very broad definition of access to treatment or evaluation to increase confidence that treatment effects due to the intervention had been captured. There is reasonable

support for instrument strength in that intervention status significantly predicts both the appropriate treatment and access to treatment/evaluation indicators in bivariate-probit analyses (see Tables A2 and A6 in Appendix); and for the access indicator intervention differences that are significant in unadjusted bivariate analyses (chi-square (1) = 6.25, $p = 0.01$).

Our indicators of appropriate treatment are broad and approximate meeting at least minimal standards for acceptable care rather than identifying what the best care strategies might accomplish. However, this study illustrates that even when selection effects obscure treatment-outcome relationships, then econometric techniques can reveal a positive effect of appropriate treatment on outcomes for diverse, depressed primary care youth. This is an important clinical and methods contribution to treatment effectiveness research concerning adolescent depression.

Clinical Significance

The present findings highlight the potential value of QI efforts designed to improve access to evidence-based treatment through primary care. Results indicate that when treatment consistent with at least minimal standards of appropriate care is received by depressed youths, but delivered largely under usual practice conditions, then it is likely to be substantially effective over 6 months, including for key underserved and minority groups. These findings are relevant to clinical practice and health policy in an era of reform and provide further empirical support beyond QI intervention findings for the clinical value of efforts to export evidence-based treatment strategies for youth depression to community practice settings.

Disclosures

Dr. Wells is an employee of the University of California, Los Angeles, and RAND corporation; receives research grants from the National Institute of Health and the Agency for Healthcare and Research Quality as well as NIH travel reimbursements for presenting at scientific meetings. Dr. Tang is an employee of the University of California, Los Angeles, and has no financial conflicts of interest to disclose. Dr. Carlson is an employee of the Stony Brook University School of Medicine; receives research grants from the National Institute of Mental Health, GlaxoSmithKline, and Bristol Myers Squibb; serves as a consultant for Eli Lilly and educational programs for the American Academy of Child and Adolescent Psychiatry; and received honoraria from Grand Rounds at various institutions. Dr. Asarnow is an employee of the University of California, Los Angeles; receives research grants from the National Institute of Mental Health; and received honoraria from the California Institute of Mental Health, Hathaways-Sycamores, Casa Pacifica, and the Melissa Institute.

Acknowledgments

The authors acknowledge the participating practice organizations and their representatives who made the Youth Partners in Care study possible: UCLA Medical Center and Venice Family Clinic (Martin Anderson, M.D., M.P.H.); Kaiser Permanente Los Angeles Medical Center (Ann P. LaBorde, Ph.D., Psy.D., Robert Zeledon, M.D.); Ventura County Medical Center (Christopher Landon, M.D.); Children's Hospital, University of Pittsburgh (Pamela J. Murray, M.D., M.P.H., David Brent, M.D., Frances Wren, M.D.); and Lisa Jaycox, Ph.D. who served as a co-PI in the original study. They thank the DSMB (Donald Guthrie, Ph.D., John Weisz, Ph.D., Mary Jane Rotheram-Borus, Ph.D.), the many colleagues and participants who contributed to the original study, and Tom Belin, Ph.D., and Susan Ettner, Ph.D. for statistical consultation.

References

Angrist JD, Imbens GW, Rubin DB: Identification of causal effects using instrumental variables. J Am Stat Assoc 91:444–455, 1996.

Asarnow JR, Carlson G, Schuster M, Miranda J, Jackson-Triche M, Wells KB: Youths Partners in Care: Clinician Guide to Depression Assessment and Management Among Youth in Primary Care Settings. Los Angeles, CA: UCLA School of Medicine; 1999. Adapted from: Rubenstein L, Unutzer J, Miranda J, et al. Partners in Care: Clinician Guide to Depression Assessment and Management in Primary Care Settings. Santa Monica, CA: RAND; 1996.

Asarnow JR, Jaycox LH, Clarke G, Lewinsohn P, Hops H, Rohde P: Stress and Your Mood: A Manual. Los Angeles, CA: UCLA School of Medicine; 1999.

Asarnow JR, Jaycox LH, Duan N, LaBorde AP, Rea MM, Murray P, Anderson M, Landon C, Tang L, Wells KB: Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: A randomized controlled trial. JAMA 293:311–319, 2005.

Asarnow JR, Jaycox LH, Tang L, Duan N, LaBorde AP, Zeledon LR, Anderson M, Murray P, Landon C, Rea M, Wells KB: Are there long-term benefits of short-term quality improvement interventions for depressed youth in primary care? Am J Psychiatry 166:1002–1010, 2009.

Barry CL, Huskamp HA: Moving beyond Parity—Mental health and addiction care under the ACA. N Engl J Med, 2011. Available at www.nejm.org/doi/full/10.1056/NEJMp1108649?ssource=hcrc

Berwick DM, Murphy JM, Goldman PA, Ware JE Jr., Barsky AI, Weinsten MC: Performance of a five-item mental health screening tests. Med Care 29:169–176, 1991.

Bhattacharya J, Goldman D, McCaffrey D: Estimating probit models with self-selected treatments. Stat Med 25:389–413, 2006.

Birmaher B, Brent D, AACAP Work Group on Quality Issues, Bernet A, Burkstein O, Walter H, Benson RS, Christman A, Farchione T, Greenhill L, Hamilton J, Keable H, Kinlan J, Schoettle U, Stock S, Ptakowsi KK, Medicus J: Practice parameter for the assessment and treatment of children and adolescents with depressive disorders. J Am Acad Child Adolesc Psychiatry 46:1503–1526, 2007.

Brent D, Emslie G, Clarke G, Wagner KD, Asarnow JR, Keller M, Vitiello B, Ritz L, Iyengar S, Abebe K, Birmaher B, Ryan N, Kennard B, Hughes C, DeBar L, McCracken J, Strober M, Suddath R, Spirito A, Leonard H, Melhem N, Porta G, Onorato M, Zelazny J: Switching to another SSRI or to venlafaxine with or without cognitive behavioral therapy for adolescents with SSRI-resistant depression: The TORDIA randomized controlled trial. JAMA 299:1683–1696, 2009.

Campo JV, Bridge JA: Treatment of youth depression. Am J Psychiatry 166:958–960, 2009.

Clarke GN, Hawkins W, Murphy M, Sheeber LB, Lewinsohn PM, Seeley JR: Targeted prevention of unipolar depressive disorder in an at-risk sample of high school adolescents. J Am Acad Child Adolesc Psychiatry 34:312–321, 1995.

Clarke GN, Lewinsohn PM, Hops H, Seeley JR: Adolescent Coping with Depression Course. Eugene, OR: Castalla Press; 1990.

Clarke GN, Rohde P, Lewinsohn PM, Hops H, Seeley J: Cognitive-behavioral treatment of adolescent depression. J Am Acad Child Adolesc Psychiatry 38:272–279, 1999.

Eaton DK, Kann L, Kinchen S, Shanklin S, Ross J, Hawkins J, Harris WA, Lowry R, McManus T, Chyen D, Lim C, Whittle L, Brener

ND, Wechsler H, Centers for Disease Control and Prevention (CDC): Youth risk behavior surveillance—United States, 2009. MMWR Surveill Summ 59:1–142, 2009.

Emslie G, Rush JA, Weinbger WA, Kowatch RA, Hughes CW, Carmody T, Rintelmann J: A double-blind, randomized placebo-controlled trial of fluoxetine in depressed children and adolescents. Arch Gen Psychiatry 54:1031–1037, 1997.

Garber J, Clarke GN, Weersing VR, Beardslee WR, Brent DA, Gladstone TR, DeBar LL, Lynch FL, D'Angelo E, Hollon SD, Shamseddeen W, Iyengar S: Prevention of depression in at-risk adolescents: A randomized controlled trial. JAMA 301:2215–2224, 2009.

Goodyer IM, Dubicka B, Wilkinson P, Kelvin R, Roberts C, Byford S, Breen S, Ford C, Barrett B, Leech A, Rothwell J, White L, Harrington R: A randomised controlled trial of cognitive behaviour therapy in adolescents with major depression treated by selective serotonin reuptake inhibitors. The ADAPT trial. Health Technol Assess 12:iii-iv, ix-60, 2008.

Heckman J: Randomization as an instrumental variable: Notes. RE Stat 78:336–341, 1996.

Horwitz SM, Hoagwood K, Stiffman AR, Summerfeld T, Weisz JR, Costello EJ, Rost K, Bean DL, Cottler L, Leaf PJ, Roper M, Norquist G: Measuring youths' use of mental health services: Reliability of the Services Assessment for Children and Adolescents (SACA). Psychiatry Serv 52:1088–1094, 2001.

Horwitz SM, Leaf PJ, Leventahl JM, Forsthy B, Speechely KN: Identification and management of psychosocial and developmental problems in community-based, primary care pediatric practices. Pediatrics 89:480–485, 1992.

Hughes CW, Emslie GJ, Crismon MO, Wagner KD, Birmaher B, Geller B, Pliszka SR, Ryan ND, Strober M, Trivedi MH, Toprac MG, Sedillo A, Llana ME, Lopez M, Rush AJ: The Texas Children's Medication Algorithm Project: Report of the Texas Consensus Conference Panel on Medication Treatment of Childhood Major Depressive Disorder. J Am Acad Child Adolesc Psychiatry 38:1442–1454, 1999.

Imbens GW, Rosenbaum PR: Randomization inference with an instrumental variable. J R Stat Soc Series A 168:109–126, 2005.

Imbens GW, Rubin DB: Bayesian inference for causal effects in randomized experiments with noncompliance. Ann Stat 25:305–327, 1997.

Jensen PS, Bhatara VS, Vitiello B, Hoagwood K, Feil M, Burke LB: Psychoactive medication prescribing practices for U.S. children. J Am Acad Child Adolesc Psychiatry 38:557–565, 1999.

Kessler RC: Epidemiology of depression. In: Handbook of Depression, edited by I. Gotlib, CL. Hammen. New York: Guildford Press; 2002; pp. 23–42.

Korn EL, Graubard BI: Predictive margins with survey data. Biometrics 55:652–659, 1999.

March J, Silva S, Petrycki S, Curry J, Wells K, Fairbank J, Burns B, Domino M, McNulty S, Vitiello B, Severe J: Treatment for Adolescents with Depression Study (TADS) Group. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression. JAMA 292:807–820, 2004.

McLellan M, McNeil BJ, Newhouse JP: Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? JAMA 272:859–866, 1994.

Miranda J, Schoenbaum M, Sherbourne C, Duan N, Wells K: Effects of Primary care depression treatment on minority patients' clinical status and employment. Arch Gen Psych 61:827–834, 2004.

Mufson L, Dorta KP, Wickramaratne P, Nomura Y, Olfson M, Weissman MM: A randomized effectiveness trial of interpersonal psychotherapy for depressed adolescents. Arch Gen Psychiatry 62:577–584, 2004.

Radloff LS: The CES-D Scale: A self report depression scale for research in the general population. Appl Psychol Meas 1:385–401, 1977.

Rosenbaum PR: Design of Observational Studies. New York: Springer; 2010.

Schoenbaum M, Unutzer J, McCaffrey D, Duan N, Sherbourne C, Wells KB: The effects of primary care depression treatment on patients' clinical status and employment. Health Services Res 37:1145–1158, 2002.

Sturm R: Instrumental variable methods for effectiveness research. Int J Methods Psychiatr Res 7:17–26, 2006.

Sturm R, Wells KB: How can care for depression become more cost-effective? JAMA 273:51–58, 1995.

U.S. Department of Health and Human Services: Mental Health: A Report of the Surgeon General. Rockville, MD: U.S. Department of Health and Human Services; 1999.

U.S. Food and Drug Administration: Public Health Advisory: Suicidality in Children and Adolescents Being Treated with Antidepressant Medications, October 15, 2004. Available at www.fda.gov/cder/drug/antidepressants/default.htm

Vitiello B, Swedo S: Antidepressant medications in children. N Engl J Med 350:1489–1491, 2004.

Wagner KD, Ambrosini P, Rynn M, Wohlberg C, Yang R, Greenbaum MS, Childress A, Donnelly C, Deas D: Sertraline Pediatric Depression Study Group: Efficacy of sertraline in the treatment of children and adolescents with major depressive disorder: Two randomized controlled trials. JAMA 290:1033–1041, 2004a.

Wagner KD, Robb AS, Finding RL, Jin J, Gutierrez MM, Heydorn WE: A randomized, placebo-controlled trial of citalopram for the treatment of major depression in children and adolescents. Am J Psychiatry 161:1079–1083, 2004b.

Weisz JR, Doss AJ, Hawley KM: Youth psychotherapy outcome research: A review and critique of the evidence base. Ann Rev Psych 56:337–363, 2005.

Wells KB, Sherbourne C, Schoenbaum M, Duan N, Meredith L, Unützer J, Miranda J, Carney MF, Rubenstein LV: Impact of disseminating quality improvement programs for depression in managed primary care: A randomized controlled trial. JAMA 283: 212–220, 2000.

World Health Organization: Composite International Diagnostic Interview (CIDI): Core Version 2.1: Interviewer's Manual. Geneva, Switzerland: World Health Organization; 1997.

Address correspondence to:
Joan Rosenbaum Asarnow, Ph.D.
University of California
UCLA Medical Plaza, Room 3310
Los Angeles 300
Los Angeles, CA 90095

E-mail: jasarnow@mednet.ucla.edu

(*Appendix follows* →)

Appendix

Table A1. Predicted Percent of Severe Depression at 6 Months, by Counseling Status[a]

	Predicted % severe depression				
	5 or fewer counseling visits	*6 or more counseling visits*	*Bivariate probit*		
	% (95% CI)	*% (95% CI)*	*Coeff*	*SE*	p-*Value*
a. Full sample ($n=341$)	43.4 (38.1 to 48.8)	10.5 (6.4 to 14.7)	−1.20	0.13	<0.001
b. Disorder subsample ($n=143$)	51.9 (35.7 to 68.1)	22.5 (−9.4 to 54.4)	−0.87	0.76	0.25
c. Symptoms only subsample ($n=198$)	36.1 (29.3 to 42.9)	4.0 (0.5 to 7.5)	−1.49	0.23	<0.001
d. Latino subsample ($n=224$)	41.8 (35.3 to 48.4)	8.6 (3.8 to 13.4)	−1.24	0.17	<0.001
e. non-Latino subsample ($n=117$)	45.8 (33.7 to 57.9)	14.8 (−2.9 to 32.5)	−1.17	0.55	0.03

Sensitivity analysis for main model by excluding medication from definition of treatment.
Severe depression: CES-D ≥ 24.
[a]Results were based on treatment coefficient from bivariate probit model, adjusted for age gender and baseline MHI-5 score. In the QI condition, 36 out of 169 youths (21.3%) had ''6 or more counseling visits'' versus 22 out tof 172 youths (12.8%) in UC.
UC = usual care; QI = quality improvement; CI = confidence interval; MHI = Mental Health Inventory 5; SE = standard error; CES-D = Center for Epidemiological Studies-Depression Scale.

Table A2. Estimated Bivariate Probit Model of Appropriate Treatment and Severe Depression Outcome among Full sample and Subgroups[a]

	Treatment equation				*Outcome equation*			
	Coefficient	*SE*	T	p-*Value*	*Coefficient*	*SE*	*t*	p-*Value*
a. Full sample (n = *341)*								
Appropriate treatment					−1.23	0.13	−9.61	<0.001
QI vs. UC	0.32	0.13	2.46	0.01				
Age	−0.04	0.04	−1.18	0.24	−0.02	0.03	−0.49	0.63
Female sex	0.12	0.20	0.60	0.55	0.46	0.17	2.64	0.01
Baseline MHI-5	−0.05	0.02	−2.93	<0.01	−0.08	0.01	−5.71	<0.000
Intercept	−1.11	0.19	−5.93	<0.001	−0.45	0.16	−2.89	<0.01
b. Disorder subsample (n = *144)*								
Appropriate treatment					−1.05	0.17	−6.25	<0.001
QI vs. UC	0.28	0.18	1.58	0.11				
Age	−0.00	0.06	−0.08	0.94	0.02	0.05	0.34	0.73
Female sex	0.49	0.29	1.72	0.09	0.47	0.25	1.87	0.06
Baseline MHI-5	−0.01	0.02	−0.48	0.63	−0.07	0.02	−3.15	<0.01
Intercept	−1.04	0.27	−3.91	<0.000	−0.34	0.22	−1.51	0.13
c. Symptoms only subsample (n = *198)*								
Appropriate treatment					−1.54	0.22	−6.89	<0.001
QI vs. UC	0.36	0.21	1.72	0.09				
Age	−0.07	0.05	−1.29	0.20	−0.03	0.04	−0.71	0.48
Female sex	−0.24	0.28	−0.85	0.40	0.49	0.25	1.98	0.05
Baseline MHI-5	−0.05	0.03	−1.84	0.07	−0.07	0.02	−3.18	0.001
Intercept	−1.15	0.28	−4.09	<0.001	−0.60	0.23	−2.58	0.01
d. Latino subsample (n = *224)*								
Appropriate treatment					−1.24	0.17	−7.42	<0.001
QI vs. UC	0.39	0.18	2.21	0.03				
Age	−0.09	0.05	−1.61	0.11	−0.07	0.04	−1.65	0.10
Female sex	0.11	0.25	0.45	0.65	0.50	0.21	2.38	0.017
Baseline MHI-5	−0.04	0.02	−1.80	0.07	−0.06	0.02	−3.53	<0.001
Intercept	−1.32	0.24	−5.43	<0.001	−0.58	0.19	−3.05	<0.01
e. non-Latino subsample (n = *118)*								
Appropriate care					−1.27	0.21	−6.02	<0.001
QI vs. UC	0.27	0.21	1.28	0.20				
Age	−0.04	0.06	−0.64	0.52	0.04	0.05	0.73	0.47
Female sex	0.05	0.35	0.13	0.90	0.15	0.32	0.47	0.64
Baseline MHI-5	−0.07	0.03	−2.46	0.01	−0.14	0.03	−4.92	<0.001
Intercept	−0.74	0.31	−2.41	0.02	−0.04	0.29	−0.14	0.89

Appropriate treatment: at least 6 psychotherapy/counseling visits or 2 months of daily use of antidepressant medication. Age was entered to the mean age of 17, and MHI-5 was centered to the mean value of 19.
[a]Table presents coefficients and SE from bivariate probit models, adjusted for age, gender, and baseline MHI-5 scores.

TABLE A3. PROBIT MODELS FOR THE EFFECT OF APPROPRIATE TREATMENT ON DEPRESSION[a]

	Bivariate probit				*Two-step probit*				*Naive probit*			
	coeff	*SE*	t	p-*Value*	*coeff*	*SE*	t	*p-Value*	*coeff*	*SE*	t	p-*Value*
a. Full sample ($n=342$)	−1.23	0.13	−9.61	<0.001	−7.34	2.53	−2.90	<0.01	0.16	0.18	0.90	0.37
b. Disorder subsample ($n=144$)	−1.05	0.17	−6.25	<0.001	−2.82	2.57	−1.10	0.27	0.35	0.24	1.45	0.15
c. Symptoms only subsample ($n=198$)	−1.54	0.22	−6.89	<0.001	−14.68	6.14	−2.39	0.02	−0.14	0.31	−0.46	0.65
d. Latino subsample ($n=224$)	−1.24	0.17	−7.42	<0.001	−3.22	2.15	−1.50	0.13	0.19	0.24	0.79	0.43
e. non-Latino subsample ($n=118$)	−1.27	0.21	−6.02	<0.001	−31.44	10.73	−2.93	<0.01	0.07	0.29	0.25	0.80

Alternative models for the main treatment indicator.

Appropriate treatment: at least 6 psychotherapy or counseling sessions or use of antidepressant medication in 6 months; severe depression: CES-D ≥ 24.

Bivariate probit estimates both equations of the IV framework simultaneously; two-step probit estimates them sequentially; and naïve probit is a standard single regression estimating the effect of treatment on outcome without accounting for unmeasured selection effects.

[a]Table presents coefficients and standard errors from three probit models, adjusted for age gender and baseline MHI-5 score.

Coeff = coefficient; IV = instrumental variables.

TABLE A4. PREDICTED PERCENT OF YOUTH WITH SEVERE DEPRESSION AT 6 MONTHS, BY ANY TREATMENT (DEFINED AS COUNSELING, ANTIDEPRESSANT MEDICATION, OR MENTAL HEALTH SPECIALTY VISIT)[a]

	Predicted % severe depression				
	No treatment	*Any treatment*	*Bivariate probit*		
	% (95% CI)	*% (95% CI)*	*Coeff*	*SE*	p-*Value*
a. Full sample ($n=342$)	50.9 (45.2 to 56.6)	15.7 (11.3 to 20.0)	−1.13	0.11	<0.0001
b. Disorder subsample ($n=144$)	61.4 (52.5 to 70.3)	25.8 (17.8 to 33.8)	−1.01	0.16	<0.0001
c. Symptoms only subsample ($n=198$)	42.8 (35.5 to 50.1)	8.2 (3.7 to 12.6)	−1.29	0.17	<0.0001
d. Latino subsample ($n=224$)	48.0 (41.0 to 55.0)	14.7 (9.4 to 20.0)	−1.08	0.14	<0.0001
e. non-Latino subsample ($n=118$)	55.2 (45.3 to 65.2)	18.3 (10.5 to 26.1)	−1.25	0.20	<0.0001

Sensitivity analysis with broad treatment indicator.

Any treatment: at least one psychotherapy or counseling sessions or use of mental health specialist, use of antidepressant medication in 6 months; severe depression: CES-D ≥ 24.

[a]Results were based on treatment coefficient from bivariate probit model, adjusted for age gender and baseline MHI-5 score. In the QI condition, 66 out of 170 youths (38.8%) received ''any treatment'' versus 45 out of 172 youths (26.2%) in UC.

TABLE A5. THREE PROBIT MODELS FOR THE EFFECT OF ANY TREATMENT (COUNSELING, ANTIDEPRESSANT MEDICATION, OR MENTAL HEALTH SPECIALTY VISIT) ON SEVERE DEPRESSION[a]

	Bivariate probit				*Two-step probit*				*Naive probit*			
	coeff	*SE*	t	p-*Value*	*coeff*	*SE*	t	p-*Value*	*coeff*	*SE*	*t*	p-*Value*
a. Full sample ($n=342$)	−1.13	0.11	−10.03	<0.001	−2.88	1.16	−2.48	0.013	0.20	0.16	1.28	0.20
b. Disorder subsample ($n=144$)	−1.01	0.16	−6.36	<0.001	−2.54	2.47	−1.03	0.30	0.37	0.22	1.67	0.09
c. Symptoms only subsample ($n=198$)	−1.29	0.17	−7.53	<0.001	−3.17	1.36	−2.32	0.02	0.00	0.23	0.01	0.99
d. Latino subsample ($n=224$)	−1.08	0.14	−7.66	<0.001	−2.10	1.39	−1.51	0.13	0.29	0.20	1.44	0.15
e. non-Latino subsample ($n=118$)	−1.25	0.20	−6.31	<0.001	−3.92	2.02	−1.94	0.05	0.05	0.27	0.20	0.84

Alternative models with sensitivity analysis using broad treatment indicator.

Any treatment: at least one psychotherapy or counseling sessions or use of mental health specialist, use of antidepressant medication in 6 months; severe depression: CES-D ≥ 24.

Bivariate probit estimates both equations of the IV framework simultaneously; two-step probit estimates them sequentially; and naïve probit is a standard single regression estimating the effect of treatment on outcome without accounting for unmeasured selection effects.

[a]Table presents coefficients and SE from three probit models, adjusted for age gender and baseline MHI-5 score.

Table A6. Estimated Bivariate Probit Model of Access to Any Treatment or Specialty Visit on Severe Depression Outcome among Full sample and Subgroups[a]

	Treatment equation				*Outcome equation*			
	Coeff	*SE*	T	p-*Value*	*Coefficient*	*SE*	t	p-*Value*
a. Full sample (n = *341)*								
Appropriate treatment					−1.13	0.11	−10.03	<0.001
QI vs. UC	0.43	0.12	3.64	<0.001				
Age	−0.09	0.03	−2.58	0.01	−0.04	0.03	−1.19	0.23
Female sex	−0.12	0.18	−0.70	0.49	0.33	0.17	1.95	0.05
Baseline MHI-5	−0.05	0.02	−3.56	<0.001	−0.08	0.01	−5.91	<0.001
Intercept	−0.58	0.17	−3.49	<0.001	−0.20	0.16	−1.25	0.21
b. Disorder subsample (n = *144)*								
Appropriate treatment					−1.01	0.16	−6.36	<0.001
QI vs. UC	0.30	0.17	1.76	0.08				
Age	−0.11	0.06	−1.92	0.05	−0.03	0.05	−0.54	0.59
Female sex	0.07	0.26	0.27	0.79	0.29	0.24	1.21	0.23
Baseline MHI-5	−0.03	0.02	−1.21	0.23	−0.08	0.02	−3.41	<0.001
Intercept	−0.37	0.24	−1.51	0.13	−0.06	0.23	−0.28	0.78
c. Symptoms only subsample (n = *198)*								
Appropriate treatment					−1.29	0.17	−7.53	<0.001
QI vs. UC	0.58	0.17	3.38	<0.001				
Age	−0.07	0.05	−1.59	0.11	−0.04	0.04	−0.88	0.38
Female sex	−0.23	0.24	−0.98	0.33	0.42	0.24	1.77	0.08
Baseline MHI-5	−0.04	0.02	−1.77	0.08	−0.07	0.02	−3.23	<0.001
Intercept	−0.82	0.24	−3.42	<0.001	−0.39	0.23	−1.70	0.09
d. Latino subsample (n = 224)								
Appropriate treatment					−1.08	0.14	−7.66	<0.001
QI vs. UC	0.41	0.15	2.70	0.01				
Age	−0.13	0.05	−2.77	0.01	−0.09	0.04	−2.23	0.03
Female sex	−0.10	0.22	−0.46	0.65	0.38	0.20	1.87	0.06
Baseline MHI-5	−0.06	0.02	−3.09	<0.001	−0.07	0.02	−3.99	<0.001
Intercept	−0.74	0.21	−3.48	<0.001	−0.35	0.19	−1.82	0.07
e. non-Latino subsample (n = 118)								
Appropriate Care					−1.25	0.20	−6.31	<0.001
QI vs. UC	0.47	0.20	2.37	0.02				
Age	−0.06	0.05	−1.18	0.24	0.03	0.05	0.49	0.62
Female sex	−0.17	0.32	−0.54	0.59	0.05	0.31	0.16	0.88
Baseline MHI-5	−0.05	0.03	−1.95	0.05	−0.13	0.03	−4.68	<0.001
Intercept	−0.33	0.29	−1.15	0.25	0.21	0.29	0.74	0.46

Sensitivity analysis, full model for ''upper limit'' access variable.

Any treatment: at least one psychotherapy or counseling sessions or use of mental health specialist, use of antidepressant medication in 6 months. Age was entered to the mean age of 17, and MHI-5 was centered to the mean value of 19.

[a]Table presents coefficients and SE from bivariate probit models, adjusted for age, gender, and baseline MHI-5 scores.