

SCREENING AND MANAGEMENT OF ADOLESCENT DEPRESSION IN A LARGE
PEDIATRIC CARE NETWORK

By

ALYSSA E. McCARTHY

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Psychology

Written under the direction of

Jami F. Young

And approved by

New Brunswick, New Jersey

October, 2019

ABSTRACT OF THE DISSERTATION

Screening and Management of Adolescent Depression in a Large Pediatric Care Network

By ALYSSA E. McCARTHY

Dissertation Director:
Jami F. Young

There is low access to adolescent depression care due to lack of identification, a shortage of mental health professionals, and personal and logistical barriers to seeking treatment. Primary care has therefore been promoted as a setting to identify and manage this prevalent and impairing condition.

Data for this study came from electronic health record (EHR) extraction in a large pediatric care network with an organizational recommendation to screen for depression at age 16 well visits using a tablet-based, EHR-integrated version of the Patient Health Questionnaire – Modified for Teens (PHQ-9-M). Analyses examined rates of screening and elevated symptoms during a two-year period. For at-risk patients (those with threshold and subthreshold scores on the PHQ-9-M), analyses explored primary care providers' (PCPs') immediate responses and follow-up care over approximately one year. In a selected subsample of at-risk patients, manual EHR chart review further assessed follow-up care and adherence to guidelines for the management of adolescent depression in primary care.

Results indicated that, across 27 practices conducting screening, 76.25% ($n=6,981$) of patients attending their age 16 well visit were screened. About one-quarter of screened patients had an elevated score (6.73% threshold; 19.23% subthreshold). On

the date of their well visit, 37.28% of at-risk patients received active follow-up by their PCP (e.g., behavioral health referral, emergency procedures for suicidality) according to EHR-extracted progress note documentation. Over one year, 75.54% of patients with threshold scores and 39.97% of patients with subthreshold scores received follow-up care (e.g., depressive disorder diagnosis, antidepressant medication) according to EHR data extraction. More detailed examination via manual EHR chart review suggested higher follow-up rates.

This study demonstrated that, using current technologies, routine adolescent depression screening is feasible across diverse primary care settings. For many adolescents, screening identified previously undetected concerns and resulted in treatment initiation. However, PCPs require additional support to manage increased adolescent depression identification through universal screening programs. Investment in primary care-based care coordination and mental health services is needed to ensure high-quality treatment and ultimately decrease the burden of adolescent depression.

Acknowledgment

I would first like to thank my Advisor and Dissertation Director, Dr. Jami Young, for her wonderful mentorship and support throughout this project and my time as a graduate student. Over the past four years, she has provided invaluable wisdom and guidance and warmly encouraged my development as a graduate student, researcher, and clinician. I am eternally grateful for her belief in my abilities, dedication to helping me succeed, and kind willingness to see me through this program. She exemplifies commitment to helping vulnerable youth and teaching the next generation of psychologists, and it is an honor to be her student.

I am also grateful for Drs. G. Terence Wilson, Brian Chu, and R. Eric Lewandowski, who have provided helpful feedback and recommendations as members of my Dissertation Committee, thereby strengthening my final project. During my time at Rutgers and NYU, I have learned a tremendous amount from each of their approaches to research focused on increasing treatment quality and accessibility, as well as their perspectives on direct clinical care. Dr. Robert Gallop has also been an extraordinary resource on advanced statistical analyses, working patiently with me and supporting my growth over the past several years. Moreover, I am indebted to Morgan Bush, who provided thoughtful, conscientious, and supportive assistance on this study's manual chart review. I am also thankful for Mark Ramos' excellent data extraction for this study as well as Elizabeth Brooks' generosity in sharing her organizational knowledge and methodological guidance.

Finally, I could not have made it to this point in my graduate studies without the loving support of my family and friends. I would like to thank my lab mates Caroline

Haimm and Carolyn Spiro for charting the path and providing me with endless help and guidance. Thank you also to my cohort mates, Arielle Linsky, Min-Jeong Yang, and Christopher Hughes, who have been a source of moral support on each step of this journey. Sean Farley, thank you for being my grounding force and biggest cheerleader every day since I first entered a psychology lab at Boston College.

Table of Contents

Abstract	ii
Acknowledgment	iv
Table of Contents	vi
List of Tables	vii
Introduction.....	1
Methods.....	12
Results.....	23
Discussion.....	36
Bibliography	56

List of Tables

Table 1. Patient demographic characteristics.....	63
Table 2. Rates of screening at CHOP primary care practices.....	65
Table 3. Rates of immediate PCP responses for patients with subthreshold and threshold PHQ-9-M scores	66
Table 4. Follow-up over one year for patients with subthreshold and threshold PHQ-9-M scores according to EHR data extraction	67
Table 5. Follow-up over one year for selected subsamples of patients with subthreshold and threshold PHQ-9-M scores according to manual EHR chart review	69
Table 6. Patient demographic characteristics as predictors of PHQ-9-M score with square root transformation.....	72
Table 7. PHQ-9-M score, suicide risk, and patient demographic characteristics as predictors of immediate PCP response	73
Table 8. PHQ-9-M score, suicide risk, and patient demographic characteristics as predictors of one-year follow-up according to EHR data extraction	75
Table 9. PHQ-9-M score, suicide risk, and patient demographic characteristics as predictors of total number of follow-up types according to manual EHR chart review.....	77

Introduction

Depression is a highly prevalent condition associated with tremendous burden on a societal level (Greenberg, Fournier, Sisitsky, Pike, & Kessler, 2015; Kessler et al., 2005). According to nationally representative surveys, approximately 8% of adolescents in the United States (US) experience major depression each year (Siu, 2016). Individuals affected by depression are at increased risk of disorder recurrence (Fergusson & Woodward, 2002; Joiner, 2000), onset of other disorders (e.g., substance use disorders, anxiety disorders), and suicidal behaviors (Fergusson & Woodward, 2002; Siu, 2016). Moreover, adolescent depression has been associated with adverse educational, occupational, and social outcomes in adulthood (Fergusson & Woodward, 2002).

Research indicates that, among adolescents, subthreshold depressive symptoms are not categorically different from major depressive disorder (MDD). Subthreshold symptoms are associated with problematic psychosocial functioning, as well as increased risk of future MDD and substance use disorders (Gotlib, Lewinsohn, & Seeley, 1995). In addition, adolescent depressive symptoms are often persistent over time, with many adolescents remaining at increased risk over several years (Garrison, Jackson, Marsteller, McKeown, & Addy, 1990; Patten, Choi, Vickers, & Pierce, 2001; Young, Mufson, & Gallop, 2010).

Taken together, these findings demonstrate the importance of addressing adolescent depression at both the symptom and disorder levels. While there are effective preventive interventions and treatments for adolescent depression (Cheung, Kozloff, & Sacks, 2013; Stice, Shaw, Bohon, Marti, & Rohde, 2009), access to care remains quite low. Estimates suggest that less than 40% of adolescents with depressive disorders

receive appropriate care (Lewandowski et al., 2016; Merikangas et al., 2011). There is a national shortage of mental health professionals (Kazdin & Blase, 2011), and specifically child and adolescent psychiatrists (Thomas & Holzer, 2006). Thus, needs cannot be met by specialists alone. Families also report personal and logistical barriers to seeking mental health treatment including fear of stigma, concerns about insurance coverage, and lack of problem recognition (Gulliver, Griffiths, & Christensen, 2010; Meredith et al., 2009).

To overcome these challenges, and in line with aims to integrate physical and mental health services, primary care has been promoted as a setting to identify and manage adolescent depression (Asarnow & Miranda, 2014). Primary care visits comprise the greatest proportion of adolescent health care visits (Ziv, Boulet, & Slap, 1999), creating increased opportunities for depression identification by primary care providers (PCPs). Furthermore, families are often comfortable discussing mental health concerns with their PCPs (Cheung et al., 2013; Kelleher & Stevens, 2009). Primary care also presents an opportunity for preventive approaches, as at-risk adolescents may be identified earlier in primary care than they would present to specialty mental health services. Moreover, interventions in primary care circumvent concerns about privacy and stigma in school settings (Gillham, Hamilton, Freres, Patton, & Gallop, 2006).

Guidelines for the Identification and Management of Adolescent Depression in Primary Care

As PCPs have been charged with an increasing role in the identification and management of adolescent depression, several organizations have developed guidelines to steer their practices. There are prominent guidelines from the American Academy of

Pediatrics (AAP; Guidelines for the Management of Adolescent Depression in Primary Care [GLAD-PC] and Bright Futures Guidelines) (Cheung et al., 2007; Cheung et al., 2018; Hagan, Shaw, & Duncan, 2017; Zuckerbrot, Cheung, et al., 2007; Zuckerbrot et al., 2018) and the US Preventive Services Task Force (USPSTF) (Siu, 2016; USPSTF, 2009, 2016) focused on identification and care for adolescent depression.

In 2007, GLAD-PC recommended depression evaluations and monitoring for adolescents who presented with emotional problems as their chief complaint or who had risk indicators for depression (e.g., previous episodes of depression, trauma) (Zuckerbrot, Cheung, et al., 2007). In 2009 and 2016, the USPSTF concluded with “moderate certainty” that screening for adolescent depression has “moderate [net benefit]” for depression detection and clinical outcomes (USPSTF, 2009, p. 1225). Thus, the USPSTF recommended routine depression screening for all 12- to 18-year-olds (USPSTF, 2009, 2016). The USPSTF’s 2016 recommendation stated that screening should be conducted when there are “adequate systems in place to ensure accurate diagnosis, effective treatment, and appropriate follow-up” (Siu, 2016, p. 2). In 2017, recommendations from Bright Futures (led by AAP) aligned with those from the USPSTF and stated that PCPs should routinely screen all patients for depression beginning at age 12 (Hagan et al., 2017). Most recently, GLAD-PC updated their recommendation to also advise annual screening for depression among all adolescents ages 12 and older using a formal self-report measure (Zuckerbrot et al., 2018). This recommendation was justified by the prevalence of adolescent depression, the disorder’s persistent adverse effects on many adolescents, the current under-identification of depression in primary care settings, and the increased rates of identification when adolescents are screened with a systematic

measure. According to GLAD-PC, there is insufficient evidence to recommend a specific measure, so providers are encouraged to select a screen that has “at least minimal validation data” and is feasible for their patients, practices, and organizations (Zuckerbrot et al., 2018, p.10). In addition to annual screening, patients with depression risk factors are recommended to receive more frequent and targeted screening (Zuckerbrot et al., 2018).

While there is agreement on the utility of universal adolescent depression screening in primary care, recommendations are somewhat less consistent on diagnostic evaluations and treatment following positive screens. When screening indicates elevated depressive symptoms, GLAD-PC recommends follow-up clinical interviews to determine the appropriate diagnosis and level of impairment (Zuckerbrot, Cheung, et al., 2007; Zuckerbrot et al., 2018). The USPSTF, on the other hand, has not provided specific recommendations on diagnostic procedures (USPSTF, 2009, 2016). Regarding treatment, GLAD-PC recommends that PCPs provide supportive interventions (e.g., psychoeducation) and care management for patients of all severity levels (Cheung et al., 2013; Cheung et al., 2018; Zuckerbrot, Cheung, et al., 2007). When symptoms are moderate-to-severe, there are complicating factors, or patients do not respond to supportive interventions alone, GLAD-PC recommends that patients receive psychotherapy and/or antidepressant medication (ADM) without a period of active support and monitoring. Consultation and referrals to mental health specialists should also be considered (Cheung et al., 2013; Cheung et al., 2007; Cheung et al., 2018). Both organizations’ guidelines emphasize the need for coordination between PCPs and mental health specialists throughout treatment (Cheung et al., 2013; Cheung et al., 2007; Cheung

et al., 2018; Siu, 2016). The USPSTF guidelines specifically advocate for collaborative care, or the use of care managers to connect patients, primary care providers, and mental health specialists (Siu, 2016). Several randomized controlled trials have documented the utility of collaborative care models for the management of adolescent depression (Asarnow et al., 2005; Asarnow et al., 2009; Richardson et al., 2014).

Barriers to Guideline Implementation in Primary Care

Naturalistic research indicates that adolescent depression care often does not adhere to the guidelines for primary care settings. There is evidence that depression frequently goes unidentified in pediatric primary care. Many PCPs lack confidence in their clinical judgment to identify depression without a screening tool (Zuckerbrot & Jensen, 2006), yet standardized screening rates are extremely low (Lewandowski et al., 2013; Lewandowski et al., 2016). Surveys conducted by the National Center for Health Statistics indicated that, in 2005 to 2010, adolescent depression screening was documented in just 0.2% of ambulatory care visits. PCPs cite practice logistics (e.g., lack of time) as important barriers to screening (Zenlea, Milliren, Mednick, & Rhodes, 2014). In turn, once depression is identified, PCPs report challenges in facilitating follow-up care. While the importance of care coordination is widely acknowledged, insufficient inter-provider communication is pervasive (Lewandowski et al., 2013). In order to prevent providers from becoming overwhelmed and dissuaded from screening, efficient and effective processes must be established (Lavigne, Feldman, & Meyers, 2016).

Efforts to Support the Identification and Management of Adolescent Depression in Primary Care

In light of these challenges, there has been significant attention directed towards enhancing the feasibility and effectiveness of adolescent depression identification and management in primary care. In an initial effort to assess the feasibility of routine adolescent depression screening, Zuckerbrot and colleagues (2007) implemented a two-stage protocol involving a self-report screen (the Columbia Depression Scale [CDS]) and self-administered, computerized diagnostic interview (the depression module of the Diagnostic Interview Schedule for Children, Version IV [DISC-IV]) in a primary care practice. Eligible participants were 13- to 17-year-old patients attending health maintenance or urgent care visits. Findings supported the feasibility of screening, with 79% of eligible patients completing the CDS over the study period. Average completion time was less than five minutes, and PCPs reported that patients expressed satisfaction with the procedure. All providers wished to continue screening after the study ended, and about two-thirds wished to continue using the DISC-IV. Importantly, most PCPs reported that they felt more comfortable addressing adolescent depression and suicidality at the conclusion of the study. Nearly three-fourths of providers estimated that referrals to mental health services increased since implementation of the protocol. This study was the first to demonstrate the achievability and perceived value of routine adolescent depression screening.

In recent years, many healthcare organizations have integrated adolescent depression screens into their electronic health records (EHRs) in order to streamline their administration, scoring, and feedback. In turn, EHR integration is intended to improve

problem identification, care coordination, and treatment quality (O'Connor et al., 2016; Sudhanthar, Thakur, Sigal, & Turner, 2015). At an academic medical center, Sudhanthar and colleagues (2015) assessed outcomes in several of these domains by administering adolescent depression screens through the EHR. In the two years preceding EHR integration of adolescent depression screens, only 10% to 16% of adolescents presenting to well visits were screened with a validated depression measure. Thus, a decision was made to administer the two-item version of the Patient Health Questionnaire (PHQ-2) to all patients ages 11 and older who were presenting to primary care for well visits. When eligible patients checked in for their visit, the nurse received a reminder to administer the PHQ-2. The measure was then automatically scored in the EHR and, if the score indicated risk, the nurse followed up with the nine-item version of the Patient Health Questionnaire (PHQ-9). The PHQ-9 was also automatically scored, and the nurse received interpretative information on its results. Twelve months into the program, screening rates increased to 82% (median rate of 52.5%), and the rate of behavioral health referrals increased by 38%. The authors concluded that automatic scoring in the EHR increased provider buy-in to routine screening, and EHR integration of screening made the practice feasible in busy primary care practices (Sudhanthar et al., 2015).

In a large health maintenance organization (HMO), Lewandowski and colleagues (2016) took the additional step of examining whether an organizational recommendation to screen for adolescent depression would affect patterns in depression screening and diagnosis over three years. Starting in 2011, the HMO recommended an adolescent version of the PHQ-9 as a screening measure in primary care. From 2010 to 2012, the number of unique adolescents who completed a PHQ-9 nearly doubled across all

departments, and it increased most markedly in pediatric primary care (i.e., from 162 to 2,283). There were large increases in positive PHQ-9 screens (i.e., from 91 to 435) and positively screened patients who went on to receive a depression diagnosis (i.e., from 43 to 134) in pediatric primary care. There was also a significant shift in the location of diagnoses from specialty mental health care to primary care, suggesting that patients were being identified earlier, or that the procedure was bringing attention to at-risk patients who may not have sought out specialty mental health care (Lewandowski et al., 2016).

The authors emphasized that this was the first naturalistic study to demonstrate a link between implementation of primary care-based adolescent depression screening and an increase in depression identification. Providers' responsiveness to an organizational recommendation was also encouraging. However, while it was not possible to determine the percentage of adolescents screened in pediatric primary care specifically, only 11% of adolescents were screened across primary and specialty mental health care departments in the final year of the study. The organization's paper-based screening method may have contributed to its lower screening rate compared to that found by Sudhanthar et al. (2015) (82%). The authors also postulated that state-mandated screening, as exists in Massachusetts, may be necessary to reach a greater proportion of patients (Hacker, Penfold, et al., 2014; Kuhlthau et al., 2011; Lewandowski et al., 2016; Romano-Clarke et al., 2014). There, screening is mandated at all well visits for publicly-insured patients under the age of 21, and practices receive reimbursement for screening (Hacker, Penfold, et al., 2014; Romano-Clarke et al., 2014). It remains unclear whether such an initiative is needed or screening rates would increase with more streamlined screening practices alone.

Beyond screening, O'Connor et al. (2016) examined patterns in follow-up care for adolescents with newly identified depressive symptoms documented in the EHR. The study took place at three healthcare organizations, including two large HMOs and a network of community health centers. The participating sites were high-performing institutions with behavioral health services, "cutting-edge care," and frequent quality improvement initiatives for adolescent behavioral health care (O'Connor et al., 2016, p. E7). The study sample included 4,612 adolescents in primary care with a score in the at-risk range (i.e., 10 or higher) on an adolescent version of the PHQ-9, a new depression diagnosis, or both within a 30-day period. The authors examined rates of documented follow-up care and/or symptom monitoring using a validated depression measure in the three months following identification.

According to the EHR, nearly two-thirds of adolescents with newly identified depressive symptoms received ADM treatment (29.39%) and/or psychotherapy (36.42%). One-third of adolescents had documented symptom monitoring within three months, while 12% of adolescents who were only identified by positive PHQ-9 score received a depressive disorder diagnosis in that time period. Of note, patients in the community health centers were significantly less likely to initiate treatment following identification than patients at the HMOs, and patients identified by elevated PHQ-9 scores alone had lower treatment rates. The authors emphasized that the quality of care in these organizations was likely to exceed that of other settings, so their results may overestimate usual rates of follow-up and treatment for adolescent depression. Given that substantial proportions of patients did not receive any follow-up care in these high-performing organizations, findings suggest a continued need for improvement in the quality of

adolescent depression care (O'Connor et al., 2016). To determine the extent of this need, it is necessary to explore these outcomes across a range of primary care practices.

Current Study

In recent years, there have been innovative strategies for promoting the identification and management of adolescent depression in primary care. Research supports the feasibility and acceptability of depression screening (Zuckerbrot, Maxon, Davies, Fisher, & Shaffer, 2007), utility of integrating screens into the EHR (Lewandowski et al., 2016; O'Connor et al., 2016; Sudhanthar et al., 2015), and effectiveness of an organizational recommendation to screen (Lewandowski et al., 2016). Across several healthcare organizations, O'Connor and colleagues (2016) found that the majority (64%) of patients initiated treatment following depression identification. However, patients with at-risk PHQ-9 scores alone were significantly less likely to initiate treatment than patients with new depression diagnoses. Follow-up care also varied significantly between different types of settings, with lower rates in community health settings.

The current study expanded upon this research by further examining the identification and initial management of adolescent depression in primary care. Findings aimed to demonstrate the effects of technological advancements in screening (i.e., EHR-integrated administration and feedback) combined with an organizational recommendation to screen across a diverse network of primary care sites. Data came from a large pediatric care network with an organizational recommendation to screen for depression at age 16 well visits using an EHR-integrated version of the Patient Health Questionnaire – Modified for Teens (PHQ-9-M) (Agency for Healthcare Research and

Quality, 2016). Given the adverse effects of adolescent depression at the symptom level (Garrison et al., 1990; Gotlib et al., 1995; Patten et al., 2001) and growing body of literature indicating positive outcomes from targeted adolescent depression prevention programs (Stice et al., 2009), this study focused on patients with both threshold and subthreshold depressive symptoms according to their PHQ-9-M screens. The specific aims of the study were as follows:

Aim 1. Assess adherence to the organizational recommendation to screen at age 16 well visits using the PHQ-9-M. Examine whether there was significant variability in adherence across primary care practices.

Aim 2. Calculate what proportions of patients screened with the PHQ-9-M had elevated depressive symptoms, as indicated by scores in the threshold and subthreshold ranges.

Aim 3. Identify PCPs' immediate responses when patients had elevated depressive symptoms according to the PHQ-9-M.

Aim 4. Assess the extent to which PCPs adhered to evaluation and treatment guidelines for adolescents with threshold screens (Cheung et al, 2018; Zuckerbrot et al., 2018) in the year following screening and explore whether there was any documented follow-up for adolescents with subthreshold screens during the one-year follow-up.

Aim 5. Examine predictors of screening, elevated depressive symptoms, and follow-up over approximately one year.

Methods

Setting

Data for this study came from patient EHRs at the Children's Hospital of Philadelphia (CHOP). CHOP is a large pediatric healthcare facility and primary care provider. CHOP has primary care practices in eastern Pennsylvania and southern New Jersey.

Participants

Full sample. The full sample included all patients who were eligible for participation in CHOP's adolescent depression screening program between September 1, 2014 and August 31, 2016. Thus, patients must have been attending their age 16 well visit (i.e., 15.5 to 16.5 years old on the visit date).

Selected subsamples. Among patients with scores in the threshold and subthreshold ranges, two random subsamples (100 patients with threshold scores; 100 patients with subthreshold scores) were selected for manual EHR chart review over approximately one year (i.e., 13 months) post-symptom identification.

Screening Procedures

As part of Pennsylvania's Quality Demonstration Grant for the Children's Health Insurance Program Reauthorization Act, primary care offices at CHOP transitioned to a fully-automated electronic screening system between 2011 and 2014. Under this electronic system, patients or their parents/caregivers were automatically prompted to complete different types of screens at specific ages (e.g., an autism screen at 18 and 24 months, a behavioral problem screen at nine years). Patients attending their age 16 well

visit were prompted to complete a depression screen with the PHQ-9-M on a tablet or at a kiosk.

The PHQ-9 is a widely-used, nine-item measure of depressive symptoms and severity (Moriarty, Gilbody, McMillan, & Manea, 2015). It was originally developed as a self-report measure for adults in primary care (Spitzer, Kroenke, Williams, & Patient Health Questionnaire Primary Care Study Group, 1999), and it has since been validated with adolescents (Richardson et al., 2010). On each of nine items, patients indicate whether they have experienced the depressive symptom (e.g., “Little interest or pleasure in doing things?”) in the past two weeks on a scale ranging from 0=*not at all* to 3=*nearly every day*.

To enhance the measure’s utility for adolescent populations, the PHQ-9-M was developed based on the PHQ-9 (Spitzer et al., 1999), Revised PHQ-9 for Adolescents (Johnson, Harris, Spitzer, & Williams, 2002), and Columbia Diagnostic Interview Schedule for Children (DISC) Depression Scale (DISC Development Group, 2000). The PHQ-9-M includes the same nine core items as the PHQ-9, with the exception of minor adjustments to increase the measure’s relevance for youth depression (i.e., inquiries about irritability, weight loss, and difficulty concentrating on schoolwork). The PHQ-9-M has four supplemental items that are not included in the total score. Two items assess severity or impairment (e.g., “In the past year have you felt depressed or sad most days, even if you felt okay sometimes?”), and two inquire about suicide risk (e.g., “Has there been a time in the past month when you have had serious thoughts about ending your life?”).

While there are no psychometric studies on the PHQ-9-M to date, researchers and stakeholders do not believe it is meaningfully different from the PHQ-9 (Lewandowski et

al., 2016; TeenScreen, 2011). On the PHQ-9, Richardson et al. (2010) found an optimal cut point of 11 for adolescents, wherein a score of 11 or higher indicates likely depression (sensitivity=89.5%; specificity=77.5%). In line with this, CHOP patients were flagged as having subthreshold depressive symptoms (“BORDERLINE”) if they had a total PHQ-9-M score of five to 10, and they were flagged as having threshold depressive symptoms (“FAILED”) if they had a total score of 11 to 27. Patients were flagged for suicide risk if they endorsed thoughts of death or self-harm in the past two weeks (i.e., answered 1=*several days* to 3=*nearly every day* on “Thoughts that you would be better off dead, or of hurting yourself in some way?”) and/or reported *yes* on at least one of the supplemental suicide risk items.

Immediate Follow-Up Procedures

When the PCP opened the patient’s chart, he/she was able to view the patient’s results on the PHQ-9-M. The PCP was able to see the patient’s response on each item, indication of suicide risk on each suicidality-related item (i.e., “SUICIDE RISK” or “Not a suicide risk”), total score, and severity range. If the patient’s score was in the threshold or subthreshold range, the PCP was prompted to document his/her follow-up action in a drop-down menu embedded in the visit’s EHR note. If the patient indicated elevated depressive symptoms and suicide risk, the PCP determined his/her action based on the PHQ-9-M score and a follow-up suicidality assessment. The suicidality assessment included five initial questions (e.g., “Have you ever wanted to kill yourself?”). If the patient answered *yes* to any of those questions, he/she was asked three additional questions about current risk (i.e., “In the past week, including today, have you had any desire to kill yourself?”; “In the past week, including today, did you have a plan to kill

yourself?"; "In the past week, including today, have you tried to kill yourself?"). If the patient answered *yes* to any of the questions on current risk, the PCP was prompted to refer to a "Suicide Crisis Resources" document for instructions on how to contact a psychiatric emergency center.

The drop-down menu options for the visit's EHR note were as follows: "Suicide risk high – Followed emergency procedures"; "Primary management of depression in primary care office"; "Referred to behavioral health provider"; "Did not refer – Already receiving behavioral health services"; and "No action needed – Reviewed questionnaire, inaccurate responses." There were also resources available in the chart for both PCPs (e.g., a template for a behavioral health referral letter) and families (e.g., psychoeducation about depression, self-help strategies, information about mental health providers). PCPs were encouraged to provide these documents to patients/families when appropriate.

Follow-up in the Year Following Identification

EHR data extraction. Thirteen months was selected as the appropriate follow-up period in order to include data from patients' subsequent well visits, since most insurers limit the frequency of well visits to once per year (Owolabi & Simpson, 2012). An information analyst used EHR data extraction to determine whether, during the follow-up period, patients had evidence of behavioral health referrals according to consult documentation or progress note keywords; new depressive disorder diagnoses according to diagnostic codes from the International Classification of Diseases, Ninth Revision, Clinical Modification or International Classification of Diseases, Tenth Revision, Clinical Modification (American Psychiatric Association, 2013); new ADM prescriptions

according to the Food and Drug Administration's (FDA's) list of ADMs with Medication Guides (FDA, 2018); and/or PHQ-9-M re-administration.

Manual EHR chart review. To further understand the types of follow-up care received, exploratory manual EHR chart reviews were completed for randomly selected subsamples of patients with threshold ($n=100$) and subthreshold ($n=100$) symptoms. The 2018 GLAD-PC recommendations on the identification, assessment, treatment, and management of adolescent depression provided a guideline on appropriate care for adolescents with depressive symptoms (Cheung et al., 2018; Zuckerbrot et al., 2018). Coders tracked the presence of behavioral health referrals, depressive disorder diagnoses, supportive monitoring, care coordination, and/or treatment during the year following symptom identification. To help distinguish whether mental health concerns were newly identified via PHQ-9-M screening, the manual chart review also tracked whether patients had documentation of a prior mental health diagnosis or current mental health treatment on the date of their index well visit.

To ensure the selected subsamples in the manual chart review represented adolescents from a range of practices, all adolescents with threshold and subthreshold screens were stratified by whether they were seen at a practice that was urban or suburban and with a higher or lower rate of Medicaid-insured patients (split at the median of 15.75%). Sampling within stratifications was in approximate proportion to their observed rates among patients with elevated PHQ-9-M scores in the full sample, maintaining balance to the best possible degree. Random samples were then taken from each stratum to yield 100 patients per subsample.

Coding for the manual chart review was conducted by two members of the research team. The research team developed a codebook including chart review steps and instructions on variable coding. To establish interrater reliability, the coders shared 20% of cases in each selected subsample (i.e., 20 randomly selected patients with threshold scores; 20 randomly selected patients with subthreshold scores). ICCs, with coder considered a random effect, indicated the coders' level of agreement on each variable (Shrout & Fleiss, 1979). Across all variables, the mean intra-class correlation coefficient (ICC)¹ of .87 (95% CI [.81, .93]) indicated good-to-excellent interrater reliability (Koo & Li, 2016). The coders reconciled coding differences via consensus, and the final dataset included their consensus codes. The codebook was updated to reflect several additional coding rules that resulted from the consensus discussions, and the shared cases were reviewed once again to ensure coding decisions matched these rules. The coders then randomly split the remaining 160 patients, each coding the charts of 40 more patients in the threshold subsample and 40 more patients in the subthreshold subsample. After independent coding, seven complex cases were reviewed by the coders together to ensure proper decisions were made.

Data Analysis

Aim 1. To determine the overall rate of adherence to CHOP's organizational recommendation to screen patients with the PHQ-9-M at their age 16 well visit, descriptive statistics were used to calculate the proportion of patients in the full sample who were screened with the PHQ-9-M, with a 95% confidence interval (CI). Practice-

¹ Feinstein and Cicchetti (1990) observed that findings indicate low agreement despite high observed agreement when marginals are imbalanced. In situations where one coder showed no variability in ratings, thereby yielding imbalanced marginals, prevalence-adjusted and bias-adjusted kappa was used (Byrt, Bishop, & Carlin, 1993).

level screening rates were also summarized descriptively with their respective proportions of patients screened and 95% CIs. The magnitude of variability between practices was assessed through an ICC (Shrout & Fleiss, 1979). The ICC was derived through a generalized linear mixed model (GLMM) for binary data, consisting of a random effect for practice and fixed effects for patient demographic characteristics (i.e., sex, minority versus majority racial/ethnic background, Medicaid versus other insurance,² urban versus suburban practice location; Ten Have, Kunselman, Pulkstenis, & Landis, 1998). To determine whether screening rates varied significantly across practices, total variability in screening rates was partitioned into within-practice variability and between-practice variability. The ICC term was the ratio of between-practice variability to total variability. A small ICC value (i.e., less than 5% [Raudenbush & Bryk, 2002]) would indicate minimal variability. Statistical significance of between-practice variability was assessed through the variance component estimate of the practice term. In GLMM, the variance component estimate is produced with a standard error and statistical significance test (Molenberghs & Verbeke, 2005).

Aim 2. To determine the rates of threshold and subthreshold depressive symptoms according to the PHQ-9-M, descriptive statistics were used to calculate the percentages of screened patients in the full sample who had PHQ-9-M scores from zero to four (normal range), five to 10 (subthreshold range), and 11 or higher (threshold range). Ninety-five percent CIs were produced for all estimates.

Aim 3. Analyses examined patterns in PCPs' immediate responses when patients endorsed elevated depressive symptoms. Descriptive statistics were used to determine the

² This variable primarily compared patients who used Medicaid versus private insurance. As indicated in the Patient Demographic Characteristics section, a small percentage of patients (1.19%) used self-pay.

percentages of patients with threshold and subthreshold symptoms who had each type of immediate PCP response (e.g., “Suicide risk high – Followed emergency procedures”; “Primary management of depression in primary care office”). Ninety-five percent CIs were produced for the percentages of patients who received each type of action.

Aim 4. EHR data extraction was conducted for all patients with threshold and subthreshold symptoms, and a manual chart review was completed on selected subsamples of at-risk patients to determine the presence of depressive disorder diagnoses and follow-up care over approximately one year (i.e., 13 months) following symptom identification. Descriptive statistics calculated the percentages of patients with each type of follow-up, and 95% CIs were derived.

Aim 5. Regression models were used to examine predictors of screening, elevated depressive symptoms, and follow-up during the year following symptom identification. GLMM assessed whether demographic characteristics predicted rates of PHQ-9-M screening, and multiple regression explored whether these characteristics predicted PHQ-9-M scores. According to a P-P Plot generated for the initial multiple regression model, the residuals were not normally distributed. A square root transformation was therefore applied to the variable for PHQ-9-M total score, and it was then approximately normally distributed.

For follow-up on at-risk screens, models examined symptom severity,³ suicide risk, and demographic characteristics as predictors. These predictors were included in (1) a multinomial logistic regression model exploring PCPs’ immediate responses to elevated

³ Models were run using the PHQ-9-M score as both a categorical (i.e., threshold versus subthreshold range) and continuous predictor with similar results. Given the higher sensitivity of continuous variables, this paper is focused on analyses including PHQ-9-M score as a continuous variable.

scores, (2) a series of binomial logistic regression models analyzing the likelihood of different follow-up types according to EHR data extraction, and (3) a negative binomial regression model⁴ assessing the number of follow-up types received according to the manual chart review.

To gain a sense of variables' collective predictability within the various models, the percent of variance explained was derived for multiple regression (R^2) and generalized linear models (pseudo R^2). These values must be compared cautiously because R^2 and pseudo R^2 are not based on the same mathematical formulation, although they are on the same scale (Nagelkerke, 1991).

Hypotheses

The small number of studies on adolescent depression screening and follow-up in primary care have indicated variability in rates of screening (Lewandowski et al., 2016; Sudhanthar et al., 2015) and follow-up care (O'Connor et al., 2016). Thus, this study was exploratory in nature.

Since CHOP had an organizational recommendation to screen at well visits (Lewandowski et al., 2016) and the primary care practices in this study were using a fully-automated electronic screening system during the study period (Agency for Healthcare Research and Quality, 2016), PHQ-9-M completion rates were predicted to be high. It was anticipated that there would be variability in screening rates across practices, since some practices may have experienced more challenges transitioning to the automated scoring system or faced technical difficulties with the tablets or kiosks. Based

⁴ Negative binomial regression was used to examine the number of follow-up types received due to over-dispersion (i.e., variance larger than the mean) in Poisson regression, as negative binomial regression is able to account for unobserved heterogeneity (Rodríguez, 2013).

on the findings in Richardson and colleagues' (2010) study on PHQ-9 scores among 13- to 17-year-olds in primary care, less than one-third of screened patients were expected to have scores in the subthreshold range (i.e., five to 10), and less than one-quarter were anticipated to have scores in the threshold range (i.e., 11 to 27). While Richardson et al. (2010) found that 32.58% and 25.34% of patients had subthreshold and threshold scores, respectively, their sample was likely to have a higher depression prevalence than CHOP primary care, as they oversampled youth with elevated PHQ-2 scores.

Among patients with at-risk screens, it was hypothesized that PCPs would more frequently indicate active follow-up plans in the EHR note at the time of the depression screen (i.e., use of emergency procedures, management in primary care, referral to behavioral health care) for patients with higher PHQ-9-M scores, and particularly those who also reported suicide risk. Over the following 13 months, rates of documented diagnoses and follow-up care were expected to be low. Linkage to mental health services is an ongoing challenge (Asarnow & Miranda, 2014; Romano-Clarke et al., 2014; Rushton, Bruckman, & Kelleher, 2002), and many PCPs feel unable to manage depression in primary care (Horwitz et al., 2015). Moreover, documentation of follow-up care may have been omitted from CHOP's EHR, since the large majority of behavioral health referrals at CHOP are for outside providers (CHOP, 2014). O'Connor and colleagues (2016) emphasized that a limitation of reliance on EHR data is the inability to distinguish between lack of follow-up care or failure to appropriately document it. Regardless, in line with O'Connor et al.'s (2016) finding that patients with more severe symptoms were more likely to receive treatment, it was hypothesized that patients with higher PHQ-9-M scores would have more documented follow-up care over one year.

Despite the risks associated with subthreshold symptoms (Garrison et al., 1990; Gotlib et al., 1995; Patten et al., 2001) and existence of effective prevention programs for such patients (Stice et al., 2009), rates of follow-up care were expected to be minimal among patients with subthreshold scores.

Results

Patient Demographic Characteristics

Full sample. The full sample was comprised of 10,713 patients who attended their age 16 well visit.⁵ The mean age was 16.00 years (standard deviation [*SD*]=0.28). As seen in Table 1, the sample was evenly split on patient sex (50.31% male; 49.69% female) and relatively diverse in terms of racial/ethnic background. For statistical analyses, patients were categorized as either White/non-Hispanic (56.92%) or from a racial/ethnic minority group (42.95%). The majority of patients had private insurance (75.50%), while about one-quarter had Medicaid (23.30%) and a small number used self-pay (1.19%). Approximately three-quarters (72.53%) of patients attended suburban practices, and the remaining quarter (27.47%) attended urban practices.

Within the full sample, there were 466 patients with PHQ-9-M scores in the threshold range and 1,331 patients with scores in the subthreshold range. These patients were also an average of 16 years old (mean age=16.01 [*SD*=.29] in the threshold score group; mean age=16.00 [*SD*=.29] in the subthreshold score group). As shown in Table 1, their demographic characteristics were somewhat different than the entire full sample, with higher proportions of patients who were female (65.54%), from racial/ethnic minority groups (51.87%), had Medicaid coverage (32.72%), and attended urban practices (36.73%).

Selected subsamples. The selected subsamples were comprised of 200 patients; 100 with threshold scores and 100 with subthreshold scores. The overall demographic

⁵ Due to implementation procedures in the initiation of CHOP's EHR-integrated screening program, a small subset of full sample patients screened with the PHQ-9-M ($n=85$, 1.21%) received their routine screen at the well visit following their index well visit. These patients were an average of 17.01 ($SD=.42$) years old on the date of their screen. They are included in all descriptive statistics on sample demographics.

and practice characteristics among patients in these subsamples are detailed in Table 1. These patients' characteristics generally reflected the full sample groups of patients with threshold and subthreshold scores, with two-proportion *z*-tests indicating no significant differences in demographic characteristics between the full sample and selected subsample groups of patients with elevated scores.

PHQ-9-M Screening Rates

Table 2 shows the PHQ-9-M screening rates across and within primary care practices. Across all primary care practices, 65.49% of patients attending their age 16 well visit were screened with the PHQ-9-M. Four practices⁶ were outliers with far lower screening rates than the other practices (i.e., ranging from no screening to 4.64%), either because they initiated the electronic screening system later or were using an alternative behavioral health measure during the study period. When these practices were excluded, the rate of screening among patients attending their age 16 well visit was 76.25%. Among the 27 practices that were conducting routine screening with the PHQ-9-M during the study period, the component estimate for the primary care practice term indicated statistically significant between-practice variability in screening rates (estimate=.05 [standard error=.01]; $z=3.36$; $p=.001$; 95% CI [.02, .07]). However, while the between-practice variance was .05, the total variance was 1.02. The ICC value of .05 indicated minimal variability, with only 4.50% of screening variance accounted for by primary care practice (Rodríguez & Elo, 2003). Thus, the statistical significance of the component estimate is likely attributable to the study's large sample size.

⁶ Practice 13, Practice 22, Practice 25, and Practice 29.

Rates of Elevated Depressive Symptoms

Among the 6,923 patients with complete PHQ-9-M screens,⁷ 74.09% ($n=5,129$; 95% CI [73.01%, 75.08%]) had a score in the normal range, 19.23% ($n=1,331$; 95% CI [18.30%, 20.15%]) had a score in the subthreshold range, and 6.73% ($n=466$; 95% CI [6.14%, 7.32%]) had a score in the threshold range. A substantial proportion of patients (8.62%; $n=597$; 95% CI [7.96%, 9.28%]) were flagged for suicide risk, the majority of whom also had total scores in the subthreshold (38.53%; $n=230$; 95% CI [34.61%, 42.44%]) or threshold range (44.89%; $n=268$; 95% CI [40.89%, 48.89%]). However, 16.58% ($n=99$; 95% CI [13.59%, 19.57%]) of patients flagged for suicide risk had scores in the normal range.

Immediate PCP Responses for Patients with Elevated Depressive Symptoms

The frequencies of immediate PCP responses for patients with threshold and subthreshold depressive symptoms are shown in Table 3. Among patients with elevated PHQ-9-M scores, 81.47% had EHR information on their PCP's immediate response according to the progress note's drop-down menu selection. PCPs most commonly indicated providing behavioral health referrals to patients with threshold scores (34.33%), while they most frequently indicated that patients with subthreshold scores did not require any follow-up action (33.36%). Overall, PCPs reported an active follow-up plan (i.e., emergency procedures for high suicide risk, behavioral health referral, or symptom management in primary care) for nearly half (47.00%) of patients with threshold scores and a substantial proportion (33.88%) of patients with subthreshold scores. About one-quarter (24.25%) of patients with threshold scores were already receiving behavioral

⁷ Ninety-three of the 7,016 full sample patients who were administered the PHQ-9-M did not fully complete it.

health services, compared to a smaller percentage of patients with subthreshold scores (16.00%).

One-Year Follow-Up for Patients with Elevated Depressive Symptoms

One-year follow-up according to EHR data extraction. For patients with elevated scores on the PHQ-9-M, EHR data extraction was used to gather information about the frequency of depressive disorder diagnoses and follow-up care over approximately one year (i.e., 13 months) following symptom identification.

Follow-up for patients with threshold PHQ-9-M scores. As shown in Table 4, three-quarters (75.54%) of patients with threshold scores had EHR-documented follow-up care within one year of symptom identification. Patients with threshold scores most frequently had one type of follow-up (38.63%), while 36.91% had more than one type. Nearly half (45.92%) had documentation of a depressive disorder diagnosis, 26.18% had an ADM prescription, and 3.43% were re-administered the PHQ-9-M. In total, 43.35% of patients with a threshold score received a behavioral health referral, either according to EHR drop-down menu documentation immediately after screening (34.33%) or based on other documentation during the follow-up period (9.01%).

Follow-up for patients with subthreshold PHQ-9-M scores. As also seen in Table 4, a smaller percentage of patients with subthreshold scores (39.97%) received follow-up care according to EHR data extraction. About one-quarter (27.42%) of patients with subthreshold scores received one type of follow-up care, and 12.55% received more than one type. While 15.85% had a depressive disorder diagnosis, 9.92% had an ADM prescription, and 2.70% were re-administered the PHQ-9-M. Altogether, 25.85% of patients with subthreshold scores received a behavioral health referral, either according to

EHR drop-down menu documentation immediately following the screen (21.71%) or another type of documentation during the follow-up period (4.13%).

One-year follow-up according to manual EHR chart review. Results from the manual EHR chart review validated EHR extraction data for patients included in the selected subsamples. Between the manual chart review and EHR extraction data, there was very good agreement on whether patients had a depressive disorder diagnosis ($\kappa=.83$; 95% CI [.74, .91]; $p<.001$) and/or ADM prescription ($\kappa=.88$; 95% CI [.80, .96]; $p<.001$) and good agreement on whether patients received a behavioral health referral ($\kappa=.63$; 95% CI [.53, .73]; $p<.001$) and/or PHQ-9-M re-administration ($\kappa=.66$; 95% CI [.22, 1.10]; $p<.001$) over the 13-month follow-up period. The lower agreement on behavioral health referral coding is likely due to EHR data extraction missing certain free-text and scanned documentation referral descriptions that were recognized by the coders, as evidenced by the manual chart review indicating that 38 more patients within the selected subsamples had referrals than EHR extraction on those patients. A post hoc review comparing patients with drop-down menu referrals according to the manual chart review but not EHR extraction revealed that EHR-extracted drop-down menu referrals were missed when they were indicated second to other selections (e.g., "Suicide risk high – Followed emergency procedures and Referred to behavioral health provider"). The lower agreement on PHQ-9-M re-administration is likely due to the low overall rate of that follow-up type.

Follow-up for patients with threshold PHQ-9-M scores. According to the manual chart review, most patients with threshold scores (91%; 95% CI [85.29%, 96.71%]) received some form of follow-up care over the 13 months after screening.

Patients received a mean of 4.73 ($SD=3.83$; range: 0-19) different types of follow-up. Results are shown in Table 5.

Behavioral health referrals. Following symptom identification, about three-quarters (74%) of patients with threshold scores received a behavioral health referral according to the manual chart review. Close to half (46%) of patients received a referral according to the embedded drop-down menu in the EHR, and 68% had other/additional documentation of a referral (i.e., progress note text, phone call, letter, order), with 40% having both types. Patients with non-drop-down menu referrals received a mean of 1.76 ($SD=1.08$, range: 1-6) referrals within the follow-up period, with 30% of patients receiving multiple referrals. Across all non-drop-down menu referrals in this selected subsample ($n=120$), about half (55.83%) were documented in progress note text, while 20.83% were by phone call, 18.33% were in a fax/letter, and 5.00% were by EHR order.

Assessment and diagnosis. Prior to their threshold PHQ-9-M score, a relatively small number of patients had a documented depressive disorder diagnosis (13%). However, 34% had a non-depressive disorder mental health diagnosis documented in their problem list. During the follow-up period, few patients (2%) had documentation of formal diagnostic assessments, but 38% received a new depressive disorder diagnosis.

Supportive monitoring and care coordination. Many patients with threshold scores received supportive care and treatment coordination. On the date of the well visit at which they were screened, 11% had a social work encounter to address their screen and coordinate care. Importantly, this type of follow-up was only possible at the five out of 31 primary care practices with embedded social work during the study period. Nearly half of patients (47%) received active support and monitoring during the follow-up period,

with an average of 2.89 ($SD=2.28$; range: 1-10) visits or phone calls focused on discussion about, monitoring of, or care coordination for depression/mental health concerns among those patients. Support and monitoring was mainly conducted by PCPs, although other primary care staff (e.g., social workers, nurses) and specialty providers sometimes administered such care. Twenty-eight percent of patients had documented psychoeducation about depression, and 18% developed a safety plan. Fewer patients received other types of supportive monitoring and care coordination (e.g., evidence-based treatment recommendations, PCP contact with mental health providers). Only one patient had the PHQ-9-M re-administered within the follow-up period, although two patients were administered another measure that assessed for depressive symptoms in primary care. Standardized depression measures were more frequently administered in non-primary care settings (e.g., the emergency department [ED], behavioral health) during the follow-up period.

Treatment. Treatment rates were high among patients with threshold scores. Forty-three percent of patients had documentation of psychotherapy during the follow-up period, with the large majority (89.28% with available information) receiving services outside of CHOP. One-third (33%) of patients with threshold scores were prescribed an ADM. Together, approximately one-quarter (23%) appeared to receive combination treatment (i.e., both psychotherapy and ADM). Patients who received ADM treatment had a mean of 1.48 ($SD=.87$; range: 1-5) different ADMs prescribed during the follow-up period, with 11 patients prescribed multiple types of ADMs. About half (54.55%) of patients taking an ADM reportedly received monitoring for adverse effects. Among patients who received psychotherapy or ADM treatment during the follow-up period,

chart review indicated that approximately half (57.14% with psychotherapy; 48.28% with ADM) initiated treatment before PHQ-9-M identification, although it is unknown whether this treatment was for depression or another mental health diagnosis. The higher rates of other prior diagnoses relative to depression (34% versus 13%) suggest many patients may have been receiving treatment for other disorders.

A subset of patients with threshold scores had acute mental health concerns during the year following symptom identification. Fourteen percent of patients presented to an ED with a psychiatric chief complaint on at least one occasion, and 9% were psychiatrically hospitalized, with an overlap of 7% having record of both (an) ED presentation(s) and inpatient hospitalization(s). Among patients who went to the ED during the follow-up period, 21.43% went on the date of their well visit, while 78.57% went at a later time.

Follow-up for patients with subthreshold PHQ-9-M scores. Over the 13-month period after screening, the majority of patients with subthreshold scores (63%; 95% CI [53.37%, 72.63%]) received some form of follow-up care according to the manual chart review. Patients received a mean of 1.98 ($SD=2.72$; range: 0-16) different types of follow-up. Findings are shown in Table 5.

Behavioral health referrals. Following symptom identification, 37% of patients with subthreshold scores received a behavioral health referral. Twenty-three percent received a referral according to the EHR note's embedded drop-down menu, and 33% had other/additional documentation of a referral (19% had both). Patients with non-drop-down menu referrals had a mean of 1.30 ($SD=.81$, range: 1-5) referrals, with 6% of patients with subthreshold scores receiving more than one referral. Across all non-drop-

down menu referrals ($n=43$), more than two-thirds (69.77%) were indicated in progress note text, while smaller percentages were by phone call (13.95%), fax/letter (11.63%), or EHR order (4.65%).

Assessment and diagnosis. A small number of patients (4%) with subthreshold scores had a depressive disorder diagnosis indicated in their EHR problem list prior to their index well visit, and 30% had a prior non-depressive disorder mental health diagnosis. In the 13 months following their subthreshold screen, only one patient had a formal diagnostic assessment according to the available documentation, but 16% received a new depressive disorder diagnosis.

Supportive monitoring and care coordination. Patients with subthreshold scores often received supportive care and treatment coordination during the follow-up period. On the date of their well visit, 2% had a social work encounter to address their screen. Approximately one-quarter (23%) of patients received active support and monitoring in the year following symptom identification. Among these patients, there was an average of 2.22 ($SD=2.34$, range: 1-12) visits or phone calls focused on depression/mental health concerns. Sixteen percent of patients had documented psychoeducation about depression, and 5% established a safety plan. Smaller numbers of patients received other types of monitoring and treatment coordination. Three patients had the PHQ-9-M re-administered within the follow-up period, and six had another depression measure administered in primary care.

Treatment. As expected, there were somewhat lower treatment rates among patients with subthreshold scores. Twenty-two percent of patients had documented psychotherapy (92.31% outside of CHOP according to available information), and 14%

received ADM treatment. Eleven percent of patients appeared to receive combination treatment, as indicated by documentation of both psychotherapy and ADM. Patients who took ADMs had a mean of 1.21 ($SD=.58$; range: 1-3) different ADMs prescribed during the follow-up period, with two patients receiving multiple types of ADMs. Among patients with an ADM prescription, 50% had monitoring for adverse effects.

Approximately two-thirds of patients who received psychotherapy (63.64%) and half of patients who received ADM treatment (46.15%) initiated this treatment prior to PHQ-9-M screening. Again, however, this prior treatment may have been for non-depressive disorder diagnoses, as suggested by the higher rate of other prior disorders compared to depression (30% versus 4%). During the follow-up period, 3% of patients presented to an ED with a psychiatric chief complaint on at least one occasion, all on a later date than their well visit. The same patients were all also psychiatrically hospitalized within one year of PHQ-9-M screening.

Outcome Predictors⁸

Predictors of PHQ-9-M screening. GLMM was used to examine whether primary care practice and demographic characteristics (i.e., patient sex, minority versus majority racial/ethnic group, Medicaid versus other insurance, urban versus suburban primary care practice) predicted patients' likelihood of PHQ-9-M screening. The overall model was not statistically significant ($\chi^2[9116]=9070.54, p=.63$), with a pseudo R^2 of .10% and no demographic variables significantly predicting screening rates.⁹

⁸ After the initial models including all predictors, parsimonious models were also run with only the statistically significant predictors. This paper reports on results from the initial models because the significance of predictors largely did not change in the parsimonious models.

⁹ The effect of primary care practice is described in the PHQ-9-M Screening Rates section.

Predictors of PHQ-9-M scores. Multiple regression explored whether patients' demographic characteristics predicted their PHQ-9-M score. Results of the model are shown in Table 6. The model significantly predicted PHQ-9-M score ($F[4, 6907]=118.28, p<.001$), with 6% of the variance in score explained by the demographic variables. Being female, being from a minority racial/ethnic background, having Medicaid, and attending an urban practice each independently predicted higher PHQ-9-M score.

Predictors of immediate PCP responses for patients with elevated depressive symptoms. Multinomial logistic regression was used to determine whether PHQ-9-M symptom severity, suicide risk, and/or patient demographic characteristics predicted type of immediate PCP response for patients with elevated screens. Results are shown in Table 7. The model significantly predicted outcomes over and above the intercept-only model ($\chi^2[24]=393.27, p<.001$), with 24% of the variance in immediate PCP response being explained by the predictors. Having a higher PHQ-9-M score and reporting suicide risk were independent predictors of requiring emergency procedures for suicidality, having symptoms managed in primary care, receiving a behavioral health referral, or already receiving behavioral health services compared to the reference group (i.e., no action needed). Female sex independently predicted greater likelihood of receiving a behavioral health referral or already receiving behavioral health services. Only female patients required emergency procedures for high suicide risk. Patients from minority racial/ethnic backgrounds and patients who attended urban practices were less likely to already have behavioral health services. Finally, patients who had Medicaid were more likely to receive behavioral health referrals according to the drop-down menu.

Predictors of one-year follow-up according to EHR data extraction. A series of binomial logistic regression models were used to examine the predictors of receiving each type of follow-up care (i.e., depressive disorder diagnosis, behavioral health referral, ADM prescription, and PHQ-9-M re-administration) during the year after screening. Results are detailed in Table 8. The models each significantly predicted outcome beyond their intercept-only model,¹⁰ with variance explained by the predictors ranging from 1% in the PHQ-9-M re-administration model to 14% in the depressive disorder diagnosis model.

According to EHR data extraction, having a higher PHQ-9-M score and endorsing suicide risk both independently predicted receiving a depressive disorder diagnosis, behavioral health referral, and/or ADM treatment within the follow-up period. Female sex predicted greater likelihood of receiving a behavioral health referral and/or antidepressant medication. Patients from minority racial/ethnic groups were significantly more likely to receive a behavioral health referral, but significantly less likely to be prescribed an ADM or have the PHQ-9-M re-administered. Attending an urban practice predicted lower likelihood of receiving a depressive disorder diagnosis or ADM, but higher likelihood of PHQ-9-M re-administration. Medicaid coverage predicted lower likelihood of having the PHQ-9-M re-administered.

Predictors of one-year follow-up according to manual EHR chart review.

Negative binomial regression was used to explore predictors of the number of different follow-up types patients received over one year across the selected subsamples of patients

¹⁰ Depressive disorder diagnosis model: $\chi^2(6)=278.88, p<.001$. Behavioral health referral model: $\chi^2(6)=88.69, p<.001$. Antidepressant medication prescription model: $\chi^2(6)=184.50, p<.001$. PHQ-9-M re-administration model: $\chi^2(3)=20.72, p=0.02$.

included in the manual chart review. The results are shown in Table 9. The model significantly improved upon the intercept model ($\chi^2[6]=62.63, p<.001$), with 5.40% of variance explained by the predictors. Higher symptom severity and endorsing suicide risk each independently predicted more types of follow-up during the year after screening. However, no demographic variables significantly predicted the number of follow-up types received.

Discussion

This study examined the feasibility and effectiveness of an adolescent depression screening initiative across a large and diverse pediatric care network. In particular, it explored the impact of leveraging current technologies (i.e., tablet-based screening, EHR-integrated feedback) and implementing an organizational recommendation to screen at age 16 well visits in primary care. Findings aimed to demonstrate whether such an initiative was associated with accurate identification and appropriate management of adolescent depression.

Screening Rates

Across 27 primary care practices participating in the PHQ-9-M screening program during the study period, more than three-quarters (76.25%) of patients attending their age 16 well visit were administered the measure. While between-practice variability in screening rates was statistically significant, it was rather small in magnitude, with only 4.5% of screening variance being attributable to primary care practice. This suggests universal screening was similarly achievable across a diverse group of practices using EHR-integrated administration. Overall, this study's screening rate was higher than the median screening rate of 53% reported by Sudhanthar et al. (2015) during their primary care sites' year-long implementation of a depression screening initiative, and only slightly lower than their final rate of 82% at year-end. Moreover, the rate was far higher than across a large HMO studied by Lewandowski and colleagues (2016), where 11% of adolescents were screened with the PHQ-9 across primary and specialty mental health care following an organizational recommendation to screen. It was remarkably similar to the rate of any behavioral health screening for zero- to 16-year-olds insured by

Massachusetts Medicaid (i.e., 76.31% of children who attended a well visit within the study period) under a state-mandated screening program (Hacker, Penfold, et al., 2014). This screening rate is promising given the increased focus on screening from professional organizations (USPSTF, 2016; Zuckerbrot et al., 2018) and screening requirements by Medicaid (Allen, 2017). Additional research is needed to establish continued feasibility of screening with broader implementation given the current recommendation to screen for depression in all adolescents ages 12 and older (Hagan et al., 2017; USPSTF, 2016; Young, Bush, & McCarthy, 2018; Zuckerbrot et al., 2018). Fortunately, the favorable screening rates in prior studies on larger age groups suggest outcomes will continue to be positive (Hacker, Penfold, et al., 2014; Sudhanthar et al., 2015).

Rates of Elevated Depressive Symptoms

Among patients attending their age 16 well visit who completed the PHQ-9-M, 25.96% had elevated scores (6.73% in the threshold range; 19.23% in the subthreshold range). As anticipated, these elevated score rates were lower than those found by Richardson et al. (2010), who oversampled patients with elevated PHQ-2 scores in their study evaluating the PHQ-9 as an adolescent depression screen. There are no other known studies examining rates of subthreshold PHQ-9 scores among adolescents. However, in a collaborative care study on adolescent depression treatment in primary care, Richardson et al. (2014) found a similar threshold rate to this study's (6.98% compared to 6.73%, respectively) using a cutoff score of 10 or higher.

This study's threshold rate, as well as the rate reported by Richardson et al. (2014), suggests a relatively small proportion of screened adolescents will have threshold scores that require follow-up. Yet screening will inevitably identify a subset of

adolescents who have significant mental health concerns and require intensive services, as demonstrated by the patients in this study who required emergency procedures, presented to the ED with psychiatric chief complaints, were psychiatrically hospitalized, and/or received a high number of documented follow-ups over one year. It is unclear whether elevated score rates would be similar across adolescence, although exploratory analyses suggest only modest variation among different age groups (see Appendix). Future efforts should explore the time required for follow-up and the demands on mental health resources under programs that screen adolescents at all ages.

Altogether, these findings stress the importance of screening, while also illustrating the need for practice-based behavioral health supports (e.g., embedded social workers, care managers) and community-based resources to ensure PCPs are not overburdened. As stated by Brent (2006), “Without the ability to care for the people who screen positive, screening will only be an exercise in frustration for everyone concerned” (p. 756). Integrated medical-behavioral primary care is a strategy receiving increased attention for adolescent depression management, with models such as collaborative care showing positive effects on depressive symptoms and remission rates (Asarnow, Rozenman, Wiblin, & Zeltzer, 2015; Richardson et al., 2014). Investment in such services is likely to meet the needs associated with increased depression identification and enhance practices’ overall quality of care.

Immediate PCP Responses for Patients with Elevated Depressive Symptoms

When patients indicated elevated depressive symptoms, PCPs were prompted to document their follow-up action/plan in a drop-down menu embedded in the visit’s EHR note. While the large majority (81.47%) of notes included drop-down menu

documentation, close to 20% were missing this documentation. Yet the manual chart review suggested many ‘non-adherent’ PCPs were not failing to describe a follow-up response/plan in their progress notes, but they preferred more flexibility in their descriptions; PCPs often replaced the drop-down menu section with more detailed/individualized information about their assessment and action steps. Regardless, it may be useful to incentivize appropriate follow-up by implementing a quality metric for plan documentation on the date of a positive screen. Such a measure is in place in the state of Oregon with positive results (Honsberger & King, 2017). In particular, it may be helpful to encourage PCPs to use the drop-down menu and augment with individualized information and/or to determine additional drop-down menu options that may preclude the need for free-text descriptions.

Among patients with drop-down menu documentation, findings suggest that screening led to many meaningful conversations and outcomes on the well visit date. The large majority (77.69%) of at-risk patients were not already receiving behavioral health services according to drop-down menu selections, suggesting screening identifies many adolescents who require support to connect with care. Nearly half (47.00%) of patients with threshold scores and a substantial proportion (33.88%) of patients with subthreshold scores received an active-follow-up plan. Patients frequently received behavioral health referrals, an action which may have involved verbal encouragement and/or instrumental support from their PCP. Other times, PCPs developed a plan to primarily manage symptoms in primary care, a decision that is appropriate when symptoms are less severe or specialty mental health services are not accessible (Cheung et al., 2018). This additionally demonstrates many PCPs’ willingness to provide adolescent depression care,

which counters concerns about their perceived lack of time, training, and confidence to address mental health problems (Horwitz et al., 2015). The screening also led to identification of acute concerns in several instances, with adolescents who endorsed high suicide risk receiving emergency procedures. This is particularly important since research has indicated that responses to PHQ-9 item 9 (i.e., “Thoughts that you would be better off dead, or of hurting yourself in some way?”) strongly predict risk of subsequent suicide attempts and moderately predict risk of subsequent suicide death (Simon et al., 2013).

Although patients already receiving behavioral health services did not require referrals, positive screens in these patients alerted PCPs to these mental health issues and encouraged conversations between PCPs and adolescents at the well visit. An important next step would be for PCPs to share the information from these screens with patients’ mental health providers to inform care, particularly since the manual chart review revealed that many patients who were already receiving behavioral health services had mental health diagnoses other than depression. Thus, some of these mental health providers may not have been aware of emerging depression. Inter-specialty communication is a known challenge in health care, which persists due to resource and EHR limitations (e.g., providers’ time pressure, insufficient reimbursement, lack of EHR interoperability) (Holmgren, Patel, & Adler-Milstein, 2017; O’Malley & Reschovsky, 2011) and suggests lower likelihood of such communication. Supports such as care management have been shown to increase inter-specialty communication (O’Malley & Reschovsky, 2011) and may help overcome these barriers to enhance care coordination for patients already engaged in behavioral health services as well as facilitate referrals for patients in need of them.

Beyond patients with active follow-up or existing treatment, the remaining subset (6.65% with threshold scores; 33.36% with subthreshold scores) did not receive follow-up due to reportedly indicating inaccurate responses on the PHQ-9-M. It is possible that these patients represented false positives, given that Richardson et al. (2010) found a PHQ-9 false positive rate of 22.5% in their adolescent sample. However, Richardson and colleagues (2010) noted that the large majority of their false positive group had mental health concerns other than MDD (e.g., “intermediate” depression, an anxiety disorder [Richardson et al., 2010, p. 5]). These patients were therefore still appropriately captured by the PHQ-9 and likely to benefit from follow-up care. The manual chart review also suggested that PCPs did not only select this option when their patients responded inaccurately, but also when their symptoms did not seem significant enough to warrant treatment. Thus, it may be incorrect to infer that the PHQ-9-M was not accurate/useful in these cases. This issue further highlights the need for appropriate drop-down menu options for describing follow-up responses. It also demonstrates the importance of educating PCPs about these patients’ elevated risk of future depression (Gotlib et al., 1995). Although such patients may not require formal depression treatment, they should receive more frequent and targeted monitoring, as well as assessments for other mental health concerns that may necessitate treatment (Zuckerbrot et al., 2018).

One-Year Follow-Up for Patients with Elevated Depressive Symptoms

According to EHR data extraction, the large majority (75.54%) of patients with threshold PHQ-9-M scores and 39.96% of patients with subthreshold scores had documented follow-up within approximately one year of symptom identification. Follow-up rates were even higher according to the manual chart review (91% with threshold

scores; 63% with subthreshold scores), since the coders were able to gather more detailed information and document additional types of care. Overall, the manual chart review validated the EHR extraction data on shared follow-up indicators (i.e., behavioral health referrals, depressive disorder diagnoses, ADM prescriptions, and PHQ-9-M re-administrations), thereby supporting the use of extraction methods in future research. Comparisons were not perfect, however, with variability between indicators (i.e., lower agreement on behavioral health referrals and PHQ-9-M re-administration) and EHR extraction underestimating follow-up care rates. This aligns with prior research comparing data from EHR extraction versus manual chart review and demonstrates the need for validation and cautious interpretation of extraction data (Kanger et al., 2014; Kern et al., 2013). If EHR extraction is going to be used moving forward, it will be important to recognize that follow-up rates may be underestimated with this methodology.

Given the small number of studies on follow-up to adolescent depression screening (Lewandowski et al., 2016; O'Connor et al., 2016), it is difficult to contextualize the rates and patterns of care found in this study. This is particularly challenging among patients with subthreshold scores, since this is the first known study to examine follow-up on subthreshold depressive symptoms identified in primary care. However, follow-up rates were generally higher than anticipated given PCPs' historical concerns about managing depression in primary care (Horwitz et al., 2015), the challenges in linking patients with mental health services (Asarnow & Miranda, 2014; Romano-Clarke et al., 2014; Rushton et al., 2002), and the fact that the majority of

mental health services occurred outside of CHOP and were therefore not captured in the EHR unless reported in progress notes or scanned documentation.

Behavioral health referrals. According to the manual EHR chart review, approximately three-fourths of patients with threshold PHQ-9-M scores and one-third of patients with subthreshold scores received a behavioral health referral within one year of symptom identification. These referral rates were higher than those indicated by EHR data extraction (43.35% threshold; 25.85% subthreshold), likely due to extraction missing certain free-text referrals in progress notes and scanned documentation. In addition, the manual chart review seemed to capture drop-down menu referrals that were not indicated by EHR extraction. Many patients received multiple referrals during the study period, and the chart review coders observed strong efforts by providers to link patients with care (e.g., repeated follow-up conversations with families).

There are no known studies examining behavioral health referral rates after the identification of adolescent depressive symptoms, but researchers have examined follow-up on broadband behavioral health measures administered in Massachusetts primary care practices (Hacker, Arsenault, et al., 2014; Romano-Clarke et al., 2014). Hacker and colleagues (2014), for instance, found a referral rate of 53.85% among commercially-insured adolescents with positive screens, which fell between this study's EHR data extraction (43.35%) and manual chart review (68%) referral rates for patients with threshold scores. In another chart review, Romano-Clarke et al. (2014) found similar referral rates of 44% to 59% following elevated behavioral health screens. Thus, referral rates were in line with other high-performing institutions (Hacker, Arsenault, et al., 2014; Romano-Clarke et al., 2014), although some patients still appeared to be overlooked. This

suggests a continued need to decrease the gap between identified concerns and follow-up. It is also necessary to ensure patient follow-through when referrals are made, as indicated by providers in this study frequently needing to make repeated referrals before patients initiated treatment.

Assessment and diagnosis. According to the manual chart review, relatively few patients had depressive disorder diagnoses documented in their problem list prior to PHQ-9-M administration, while about one-third had non-depressive disorder mental health diagnoses listed before screening. This demonstrates that the PHQ-9-M frequently identified patients with known mental health concerns, but likely unrecognized depression. The PHQ-9-M's validity is also supported by the frequency of new depressive disorder diagnoses during the follow-up period according to both EHR data extraction (23.65%) and manual chart review (27%). These diagnosis rates were higher than the rate reported by O'Connor and colleagues (2016) during a three-month follow-up period tracking patients identified by threshold PHQ-9 scores (12%). Given the comparatively low rates of diagnostic assessments, however, it is possible that diagnoses were provided based on clinical impressions or positive screens alone (O'Connor et al., 2016). This is concerning for treatment quality and diagnostic validity, and it should be explored further.

Importantly, the manual chart review elucidated differences in how mental health concerns/diagnoses were documented in the EHR problem list. For many patients, PHQ-9-M screening appeared helpful in prompting providers to formally indicate mental health concerns/diagnoses in the problem list. However, other patients had these issues noted in their progress note text but not their problem list. This is potentially problematic

for symptom monitoring across visits and may compromise treatment coordination among providers. For instance, there were several patients with elevated PHQ-9-M scores but without problem list documentation who repeatedly presented for medical evaluations (e.g., to cardiology for chest pain, to gastroenterology for stomach problems) during the follow-up period without significant medical findings. If mental health concerns/diagnoses had been documented more clearly, specialty providers may have more accurately/efficiently developed a holistic clinical picture. These observations suggest that providers may benefit from clearer policies on EHR documentation of positive screens, as well as mental health problems more broadly. Communication between providers at separate institutions is also essential, although there are additional barriers when the EHR is not shared (O'Malley & Reschovsky, 2011), and EHR interoperability remains limited in US hospitals (Holmgren et al., 2017).

Supportive monitoring and care coordination. Many patients received excellent supportive care and treatment coordination following symptom identification. The manual chart review demonstrated that screening often prompted impactful conversations with families, and parents/caregivers frequently stated they were previously unaware of their child's symptoms. On the date of their well visit, a subset of patients in the manual chart review (11% with threshold scores; 2% with subthreshold scores) met with a social worker to further discuss their elevated screen. The coders observed that social workers were able to promptly gather more comprehensive information from patients, thoroughly assess their treatment needs, and address potential treatment barriers. Social workers seemed particularly helpful to patients with significant symptoms and families with complex challenges. When there was not a social worker

available to help such families, PCPs often spent extensive amounts of time trying to coordinate services and provide support. These observations suggest that embedded social work was strongly beneficial to adolescent depression management in these primary care practices.

During the follow-up period, nearly half of patients with threshold scores and approximately one-quarter of patients with subthreshold scores had visits or phone calls focused on active support and monitoring of mental health symptoms. These conversations involved activities such as assessing symptoms, providing psychoeducation, coaching parents on symptom management, and encouraging treatment engagement. Substantial numbers of patients also had documented psychoeducation about depression and/or development of a safety plan. Fewer patients received other forms of supportive care (e.g., treatment planning, PHQ-9-M re-administration). These lower rates represent a departure from the GLAD-PC guidelines for adolescent depression (Cheung et al., 2018; Zuckerbrot et al., 2018), suggesting the need for additional resources to make these interventions feasible (Richardson et al., 2014), education about expected procedures (Kerker et al., 2015), and/or quality metrics to incentivize important components of the adolescent depression care pathway (Lewandowski et al., 2013). Collaborative care is a particularly promising solution, given that care managers in research studies have been able to provide a range of supportive services including psychoeducation, safety assessments, motivational interviewing, brief psychotherapy, and symptom monitoring (Richardson et al., 2014). Timely and consistent access to such services could be expected to support guideline adherence, decrease strain on PCPs, and improve care.

Treatment. Lastly, many patients had documented treatment during the follow-up period. Patients with elevated depressive symptoms frequently received ADM treatment (14.13% according to EHR data extraction; 23.50% according to the manual chart review) and/or psychotherapy (32.50% according to the manual chart review). The manual chart review suggested that approximately half of these patients initiated their treatment after PHQ-9-M symptom identification, suggesting that screening prompted treatment. Among patients already in treatment, the higher rate of other prior mental health diagnoses relative to depression (32% versus 8.50%) suggests many patients were receiving care focused on other concerns. For this subset of patients to receive appropriate treatment, it is essential that PCPs inform existing providers about the identification of elevated depressive symptoms.

In a related study, O'Connor and colleagues (2016) assessed rates of treatment initiation following initial depressive symptom identification, and findings indicated similar rates of ADM treatment (29.39%) and psychotherapy (36.42%) compared to patients with threshold scores in the current study. These results must be considered in light of O'Connor et al. (2016) using a three- rather than 13-month follow-up period and only including patients with newly identified symptoms. Nonetheless, it is impressive that the current study's rates were comparable since O'Connor and colleagues (2016) examined patients identified by PHQ-9 scores and/or depressive disorder diagnoses at "cutting-edge" institutions with frequent quality improvement initiatives for adolescent behavioral health care (O'Connor et al., 2016, p. E7).

Certain patients identified by PHQ-9-M screening had significant clinical severity, as evidenced by ED presentations with psychiatric chief complaints and/or

inpatient psychiatric hospitalizations during the follow-up period. The large majority of ED presentations (82.35%) were not on the date of the patient's well visit, and some patients had not received mental health interventions prior to these events. While more common among patients with threshold scores, several patients with subthreshold scores also required such levels of care. These findings support a focus on subthreshold in addition to threshold PHQ-9-M screens and underscore the importance of early and persistent efforts to treat existing and emerging symptoms.

Outcome Predictors

Predictors of PHQ-9-M screening. No demographic variables significantly predicted whether patients received PHQ-9-M screening. This is an encouraging finding that suggests similar accessibility of the screening program across various patient populations and a wide variety of practices.

Predictors of PHQ-9-M scores. Demographic characteristics that are typically associated with greater adolescent depression risk predicted higher PHQ-9-M scores in this study. In line with prior research indicating higher rates of depressive disorders among adolescent girls compared to boys (Hankin et al., 1998), female sex predicted higher PHQ-9-M score. Belonging to a racial/ethnic minority group also predicted more significant symptoms, which is expected given research indicating higher rates of depression among ethnic minority youth in the US, possibly related to factors including greater likelihood of discrimination (Anderson & Mayes, 2010). In addition, Medicaid coverage independently predicted higher PHQ-9-M score, which matches with previous research suggesting that youth from lower-income families have increased likelihood of developing depressive symptoms (Melchior et al., 2010). Lastly, attending an urban

practice predicted higher scores. Research indicates that youth living in urban areas tend to have increased exposure to stressors and violence (Fitzpatrick, Piko, Wright, & LaGory, 2005), and these experiences may be associated with greater risk of adolescent depression (Fitzpatrick, 1993).

Predictors of follow-up for patients with elevated depressive symptoms. As expected, higher symptom severity and suicide risk each independently predicted more active PCP responses and follow-up care over one year as assessed by EHR extraction and manual chart review. This finding demonstrates providers' adherence to guidelines on adolescent depression management, which recommend more immediate and comprehensive services for patients with more significant symptoms (Cheung et al., 2018; Zuckerbrot et al., 2018). Several demographic characteristics also predicted differences in follow-up according to EHR extraction when controlling for symptom severity and suicide risk. Demographic variables did not independently predict follow-up rates according to the manual chart review, however, likely to due to reduced statistical power in the manual chart review analyses.

According to EHR data extraction, female sex predicted greater likelihood of having preexisting behavioral health services, receiving a behavioral health referral, and/or being prescribed an ADM. These results align with previous research demonstrating that females are more likely to seek mental health care (Wang et al., 2005), potentially due to better conscious problem recognition (Kessler, Brown, & Broman, 1981).

On the other hand, patients from minority racial/ethnic backgrounds were less likely to have prior behavioral health services compared to White/non-Hispanic patients.

EHR extraction suggested that patients from minority groups were more likely to receive behavioral health referrals but less likely to receive ADM treatment and/or PHQ-9-M re-administration during the follow-up period. The lower treatment rates prior to/following screening are unsurprising, as research consistently demonstrates reduced access to behavioral health services among individuals from minority racial/ethnic groups (Garland et al., 2005; McGuire & Miranda, 2008). In particular, studies have indicated lower likelihood of ADM treatment among individuals from various minority racial/ethnic groups (Zito et al., 2003; Zito, Safer, Zuckerman, Gardner, & Soeken, 2005). Researchers have found that African-American parents tend to have less favorable attitudes towards their children's use of ADMs (Stevens et al., 2009) and that African-American and Hispanic individuals are more likely to have negative beliefs about ADMs, such as concerns about addictiveness and lack of effectiveness (Cooper et al., 2003). Of note, rates of psychotherapy were not assessed via EHR data extraction. It is therefore unclear whether the same differences in treatment rates between minority and non-minority adolescents would have been evident for non-pharmacological approaches. While it is positive that behavioral health referrals were more likely among adolescents from minority groups (potentially due to lower likelihood of preexisting care), and racial/ethnic background did not predict total types of follow-up in the manual chart review, research must continue to examine whether diverse adolescents are receiving equitable depression treatment.

Among indicators of socioeconomic status (SES), Medicaid versus private insurance coverage predicted higher likelihood of receiving a behavioral health referral according to drop-down menu selections on the well visit date, but it was not a significant

predictor of referrals over the entire follow-up period according to EHR extraction. EHR extraction may be a more reliable indicator of referral trends than immediate drop-down menu selection due to its examination of the entire follow-up period and the proportion of patients missing drop-down menu responses (18.14%). Medicaid coverage predicted lower likelihood of PHQ-9-M re-administration, which is curious in light of Medicaid's reimbursement for behavioral health screening (AAP, 2015) in contrast with challenges with reimbursement from certain private insurers (Gold, 2017; Suicide Prevention Resource Center, 2014). Yet this finding must be interpreted with caution due to the low overall rate of this follow-up type.

Attending an urban rather than suburban practice predicted lower likelihood of having prior behavioral health services, receiving a depressive disorder diagnosis during the follow-up period, and/or being treated with an ADM over one year according to EHR extraction. These are interesting results given the non-significant differences on these outcomes for Medicaid- versus privately-insured patients. It appears that factors other than SES, such as demands on urban mental health resources or logistical barriers associated with urban residence (e.g., family not having a car) or urban practice location (e.g., more patients seen), may have contributed to certain types of service access in this study's population. Also in juxtaposition to Medicaid coverage effects, patients who attended urban practices were more likely to have the PHQ-9-M re-administered during the follow-up period. The first four practices to initiate CHOP's adolescent depression screening program were urban, suggesting the possibility of more advanced incorporation of PHQ-9-M administration and follow-up during the study period. Again, however, this

finding must be interpreted carefully, since PHQ-9-M re-administration rates were very low in this study's sample.

Taken together, these findings suggest lower treatment rates among patients from certain traditionally underserved populations despite similar or higher rates of behavioral health referrals. These results elucidate the continued need for innovative and tailored approaches to linking adolescents and families with treatment and promoting their continued engagement with services. Primary care provides an opportunity for such interventions due to its broader accessibility as a location to discuss mental health concerns (Kelleher & Stevens, 2009; Mulvihill et al., 2005; Murry, Heflinger, Suiter, & Brody, 2011). For instance, a primary care-based quality improvement intervention with modest modifications for minority patients (e.g., including providers with expertise in working with minority patients on treatment teams, translating materials into different languages) led to decreased racial/ethnic disparities in depression care (Miranda et al., 2003). Such accommodations are essential first steps in providing just and equitable services across diverse patient populations.

Study Limitations

Although methodologically rigorous, this study had several limitations. Because the screening program initially only included patients attending their age 16 well visit, analyses were confined to that population for examining the effects of universal screening. Further research should assess patterns of screening and follow-up among adolescents who were attending non-age 16 well visits during the same time period (see Appendix for exploratory table) and explore whether this study's findings generalize to

all adolescents now that the screening program has expanded to youth ages 12 and older (Young et al., 2018).

Furthermore, receipt of care cannot necessarily be attributed to screening. Although EHR data extraction only included depressive disorder diagnoses and ADM prescriptions that were entered during the follow-up period, this also would have captured diagnoses that were adjusted or reentered, prescriptions that were reordered, and diagnoses or prescriptions that were historical (e.g., given by an outside provider before the index well visit but first entered in the EHR on or after the well visit date). According to EHR drop-down menu selections, about one-fifth of patients were already receiving behavioral health services on the date of their well visit. The manual chart review also indicated many patients were in treatment prior to screening.

Data also likely reflects an underestimation of follow-up/treatment rates. As emphasized by O'Connor and colleagues (2016), a limitation of reliance on EHR data is the inability to distinguish between lack of follow-up care or failure to document it. Beyond scanned documentation from other health care providers, which was read by the manual chart review coders when possible, data was confined to CHOP. This is significant since patients frequently received treatment in community clinics, schools, private practice settings, and other hospitals. Moreover, CHOP behavioral health care notes have special protections in the EHR, and their content could not be accessed for this study. While the coders were able to use billing codes to identify psychotherapy and medication monitoring visits in CHOP behavioral health care, other types of follow-up (e.g., psychoeducation, treatment planning, safety plan development) likely occurred but were unable to be coded. Although there was overall good agreement between data from

EHR extraction and the manual chart review, EHR extraction less effectively identified behavioral health referrals, as evidenced by the higher referral rate according to the manual chart review and indication that EHR extraction missed certain drop-down menu referrals by comparing discrepancies. Due to the variable accuracy of extracted data (Kern et al., 2013), it is necessary to interpret extraction findings with caution. The manual chart review also tracked additional follow-up types that were not feasible to include in the EHR extraction, further demonstrating the constraints of this powerful data collection method.

Lastly, this study did not compare rates of depression care to a time period before the screening program started due to its focus on follow-up for patients identified by universal screening. A future study might examine rates of documented behavioral health referrals, depressive disorder diagnoses, and depression treatment among all adolescents seen at primary care practices before and after screening initiation. Such work could build upon the findings of Lewandowski et al. (2016) and clarify the overall impact of routine screening on rates of adolescent depression identification and treatment.

Conclusions

This study demonstrates that adolescent depression can be reliably identified using a brief measure in primary care, and screening leads to treatment initiation for many adolescents. However, PCPs require additional support to connect adolescents with timely, high-quality services, particularly as adolescents receive more frequent screening in adherence with recent guidelines (Hagan et al., 2017; USPSTF, 2016; Zuckerbrot et al., 2018). Integrated care models have shown promise for adolescent depression outcomes (Asarnow et al., 2015), and they may be especially instrumental in reducing

care disparities for underserved populations (Miranda et al., 2003). For adolescents who do not require formal treatment, integrated care may involve supportive monitoring or preventive interventions to protect against the onset of more significant depression (Stice et al., 2009). As universal screening for adolescent depression becomes more widespread, it is essential to dedicate resources to help PCPs link identified adolescents with appropriate services to ensure adolescents' safety, well-being, and long-term health.

Bibliography

- AAP. (2015). Medicaid reimbursement survey, 2015: Pennsylvania. Elk Grove Village, IL: American Academy of Pediatrics.
- Agency for Healthcare Research and Quality. (2016). Introducing electronic screening tools for developmental delay and autism into pediatric primary care *The national evaluation of the CHIPRA quality demonstration grant program*. Philadelphia, PA: PolicyLab at The Children's Hospital of Philadelphia.
- Allen, L. M. (2017). *Medical assistance bulletin*. Retrieved from http://www.dhs.pa.gov/cs/groups/webcontent/documents/bulletin_admin/c_264244.pdf.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Arlington, VA: American Psychiatric Publishing.
- Anderson, E. R., & Mayes, L. C. (2010). Race/ethnicity and internalizing disorders in youth: A review. *Clinical Psychology Review, 30*(3), 338-348.
- Asarnow, J. R., Jaycox, L. H., Duan, N., LaBorde, A. P., Rea, M. M., Murray, P., . . . Wells, K. B. (2005). Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: A randomized controlled trial. *Journal of the American Medical Association, 293*(3), 311-319.
- Asarnow, J. R., Jaycox, L. H., Tang, L., Duan, N., LaBorde, A. P., Zeledon, L. R., . . . Wells, K. B. (2009). Long-term benefits of short-term quality improvement interventions for depressed youths in primary care. *American Journal of Psychiatry, 166*(9), 1002-1010.
- Asarnow, J. R., & Miranda, J. (2014). Improving care for depression and suicide risk in adolescents: Innovative strategies for bringing treatments to community settings. *Annual Review of Clinical Psychology, 10*, 275-303.
- Asarnow, J. R., Rozenman, M., Wiblin, J., & Zeltzer, L. (2015). Integrated medical-behavioral care compared with usual primary care for child and adolescent behavioral health: A meta-analysis. *JAMA Pediatrics, 169*(10), 929-937.
- Brent, D. A. (2006). Screens and doors: The management of adolescent depression in primary care. *Archives of Pediatrics and Adolescent Medicine, 160*(7), 755-756.
- Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence, and kappa. *Journal of Clinical Epidemiology, 46*(5), 423-429.
- Cheung, A. H., Kozloff, N., & Sacks, D. (2013). Pediatric depression: An evidence-based update on treatment interventions. *Current Psychiatry Reports, 15*(8), 381.
- Cheung, A. H., Zuckerbrot, R. A., Jensen, P. S., Ghalib, K., Laraque, D., Stein, R. E., & Glad-PC Steering Group. (2007). Guidelines for Adolescent Depression in Primary Care (GLAD-PC): II. Treatment and ongoing management. *Pediatrics, 120*(5), e1313-e1326.
- Cheung, A. H., Zuckerbrot, R. A., Jensen, P. S., Laraque, D., Stein, R. E. K., & Glad-PC Steering Group. (2018). Guidelines for Adolescent Depression in Primary Care (GLAD-PC): Part II. Treatment and ongoing management. *Pediatrics, 141*(3), e20174082.
- CHOP. (2014). Caring for your child: Screening for teen stress or depression: Next steps. Philadelphia, PA: The Children's Hospital of Philadelphia.

- Cooper, L. A., Gonzales, J. J., Gallo, J. J., Rost, K. M., Meredith, L. S., Rubenstein, L. V., . . . Ford, D. E. (2003). The acceptability of treatment for depression among African-American, Hispanic, and White primary care patients. *Medical Care*, *41*(4), 479-489.
- DISC Development Group. (2000). Interviewer manual: User manual. Joy and William Ruane Center to Identify and Treat Mood Disorders: Division of Child Psychiatry, Columbia University.
- FDA. (2018). List of antidepressant drugs with medication guides. Retrieved May 23, 2018, from <https://www.fda.gov/downloads/Drugs/DrugSafety/InformationbyDrugClass/UCM161647.pdf>
- 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.
- Fergusson, D. M., & Woodward, L. J. (2002). Mental health, educational, and social role outcomes of adolescents with depression. *Archives of General Psychiatry*, *59*(3), 225-231.
- Fitzpatrick, K. M. (1993). Exposure to violence and presence of depression among low-income, African-American youth. *Journal of Consulting and Clinical Psychology*, *61*(3), 528-531.
- Fitzpatrick, K. M., Piko, B. F., Wright, D. R., & LaGory, M. (2005). Depressive symptomatology, exposure to violence, and the role of social capital among African American adolescents. *American Journal of Orthopsychiatry*, *75*(2), 262-274.
- Garland, A. F., Lau, A. S., Yeh, M., McCabe, K. M., Hough, R. L., & Landsverk, J. A. (2005). Racial and ethnic differences in utilization of mental health services among high-risk youths. *American Journal of Psychiatry*, *162*(7), 1336-1343.
- Garrison, C. Z., Jackson, K. L., Marsteller, F., McKeown, R., & Addy, C. (1990). A longitudinal study of depressive symptomatology in young adolescents. *Journal of the American Academic of Child & Adolescent Psychiatry*, *29*(4), 581-585.
- Gillham, J. E., Hamilton, J., Freres, D. R., Patton, K., & Gallop, R. (2006). Preventing depression among early adolescents in the primary care setting: A randomized controlled study of the Penn Resiliency Program. *Journal of Abnormal Child Psychology*, *34*(2), 203-219.
- Gold, J. (2017). Health insurers are still skimping on mental health coverage. *Treatments*. Retrieved May 28, 2018, from <https://www.npr.org/sections/health-shots/2017/11/29/567264925/health-insurers-are-still-skimping-on-mental-health-coverage>
- Gotlib, I. H., Lewinsohn, P. M., & Seeley, J. R. (1995). Symptoms versus a diagnosis of depression: Difference in psychosocial functioning. *Journal of Consulting and Clinical Psychology*, *63*(1), 90-100.
- Greenberg, P. E., Fournier, A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *Journal of Clinical Psychiatry*, *76*(2), 155-162.
- Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry*, *10*(1), 113.

- Hacker, K. A., Arsenaault, L., Franco, I., Shaligram, D., Sidor, M., Olfson, M., & Goldstein, J. (2014). Referral and follow-up after mental health screening in commercially insured adolescents. *Journal of Adolescent Health, 55*(1), 17-23.
- Hacker, K. A., Penfold, R., Arsenaault, L., Zhang, F., Murphy, J. M., & Wissow, L. (2014). Screening for behavioral health issues in children enrolled in Massachusetts Medicaid. *Pediatrics, 133*(1), 46-54.
- Hagan, J. F., Shaw, J. S., & Duncan, P. M. (2017). *Bright Futures: Guidelines for health supervision of infants, children, and adolescents* (4th Ed.). Elk Grove Village, IL: American Academy of Pediatrics.
- Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology, 107*(1), 128-140.
- Heinzel, H., & Mittlböck, M. (2003). Pseudo R-squared measures for Poisson regression models with over- or underdispersion. *Computational Statistics & Data Analysis, 44*(1-2), 253-271.
- Holmgren, A. J., Patel, V., & Adler-Milstein, J. (2017). Progress in interoperability: Measuring US hospitals' engagement in sharing patient data. *Health Affairs, 36*(10), 1820-1827.
- Honsberger, K., & King, A. (2017). State strategies to measure and incentivize adolescent depression screening and treatment in Medicaid. Washington, DC: National Academy for State Health Policy.
- Horwitz, S. M., Storfer-Isser, A., Kerker, B. D., Szilagyi, M., Garner, A., O'Connor, K. G., . . . Stein, R. E. K. (2015). Barriers to the identification and management of psychosocial problems: Changes from 2004 to 2013. *Academic Pediatrics, 15*(6), 613-620.
- Johnson, J. G., Harris, E. S., Spitzer, R. L., & Williams, J. B. (2002). The Patient Health Questionnaire for Adolescents: Validation of an instrument for the assessment of mental disorders among adolescent primary care patients. *Journal of Adolescent Health, 30*(3), 196-204.
- Joiner, T. E. (2000). Depression's vicious scree: Self-propagating and erosive processes in depression chronicity. *Clinical Psychology: Science and Practice, 7*(2), 203-218.
- Kanger, C., Brown, L., Mukherjee, S., Xin, H., Diana, M. L., & Khurshid, A. (2014). Evaluating the reliability of EHR-generated clinical outcomes reports: A case study. *EGEMS (Washington, DC), 2*(3).
- Kazdin, A. E., & Blase, S. L. (2011). Booting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science, 6*(1), 21-37.
- Kelleher, K. J., & Stevens, J. (2009). Evolution of child mental health services in primary care. *Academic Pediatrics, 9*(1), 7-14.
- Kerker, B. D., Chor, K. H. B., Hoagwood, K., Radigan, M., Perkins, M., Setais, J., . . . Horwitz, S. M. (2015). Detection and treatment of mental health issues by pediatric PCPs in New York State: An evaluation of Project TEACH. *Psychiatric Services, 66*(4), 430-433.

- Kern, L. M., Malhotra, S., Barrón, Y., Quaresimo, J., Dhopeswarkar, R., Pichardo, M., . . . Kaushal, R. (2013). Accuracy of electronically reported “meaningful use” clinical quality measures: A cross-sectional study. *Annals of Internal Medicine*, *158*(2), 77-83.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 593-602.
- Kessler, R. C., Brown, R. L., & Broman, C. L. (1981). Sex differences in psychiatric help-seeking: Evidence from four large-scale surveys. *Journal of Health and Social Behavior*, *22*(1), 49-64.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155-163.
- Kuhlthau, K., Jellinek, M. S., White, G., Vancleave, J., Simons, J., & Murphy, J. M. (2011). Increases in behavioral health screening in pediatric care for Massachusetts Medicaid patients. *Archives of Pediatrics and Adolescent Medicine*, *165*(7), 660-664.
- Lavigne, J. V., Feldman, M., & Meyers, K. M. (2016). Screening for mental health problems: Addressing the base rate fallacy for a sustainable screening program in integrated primary care. *Journal of Pediatric Psychology*, *41*(10), 1081-1090.
- Lewandowski, R. E., Acri, M. C., Hoagwood, K. E., Olfson, M., Clarke, G., Gardner, W., . . . Horwitz, S. M. (2013). Evidence for the management of adolescent depression. *Pediatrics*, *132*(4), e996-e1009.
- Lewandowski, R. E., O'Connor, B., Bertagnoli, A., Beck, A., Tinoco, A., Gardner, W. P., . . . Horwitz, S. M. (2016). Screening for and diagnosis of depression among adolescents in a large health maintenance organization. *Psychiatric Services*, *67*(6), 636-641.
- McGuire, T. G., & Miranda, J. (2008). New evidence regarding racial and ethnic disparities in mental health: Policy implications. *Health Affairs (Millwood)*, *27*(2), 393-403.
- Melchior, M., Chastang, J. F., Walburg, V., Arseneault, L., Galera, C., & Fombonne, E. (2010). Family income and youths' symptoms of depression and anxiety: A longitudinal study of the French GAZEL Youth cohort. *Depression and Anxiety*, *27*(12), 1095-1103.
- Meredith, L. S., Stein, B. D., Paddock, S. M., Jaycox, L. H., Quinn, V. P., Chandra, A., & Burnam, A. (2009). Perceived barriers to treatment for adolescent depression. *Medical Care*, *47*(6), 677-685.
- Merikangas, K. R., He, J. P., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., . . . Olfson, M. (2011). Service utilization for lifetime mental disorders in U.S. adolescents: Results of the National Comorbidity Survey-Adolescent Supplement (NCS-A). *Journal of the American Academic of Child & Adolescent Psychiatry*, *50*(1), 32-45.
- Miranda, J., Duan, N., Sherbourne, C., Schoenbaum, M., Lagomasino, I., Jackson-Triche, M., & Wells, K. B. (2003). Improving care for minorities: can quality improvement interventions improve care and outcomes for depressed minorities?

- Results of a randomized, controlled trial. *Health Services Research*, 38(2), 613-630.
- Molenberghs, G., & Verbeke, G. (2005). *Models for discrete longitudinal data*. New York, NY: Springer-Verlag New York.
- Moriarty, A. S., Gilbody, S., McMillan, D., & Manea, L. (2015). Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): A meta-analysis. *General Hospital Psychiatry*, 37(6), 567-576.
- Mulvihill, B. A., Jackson, A. J., Mulvihill, F. X., Romaine, M., Gyaben, S., Telfair, J., & Caldwell, C. (2005). The impact of SCHIP enrollment on adolescent-provider communication. *Journal of Adolescent Health*, 37(2), 94-102.
- Murry, V. M., Heflinger, C. A., Suiter, S. V., & Brody, G. H. (2011). Examining perceptions about mental health care and help-seeking among rural African American families of adolescents. *Journal of Youth & Adolescence*, 40(9), 1118-1131.
- Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3), 691-692.
- O'Connor, B. C., Lewandowski, R. E., Rodriguez, S., Tinoco, A., Gardner, W., Hoagwood, K., & Scholle, S. H. (2016). Usual care for adolescent depression from symptom identification through treatment initiation. *JAMA Pediatrics*, 170(4), 373-380.
- O'Malley, A. S., & Reschovsky, J. D. (2011). Referral and consultation communication between primary care and specialist physicians: Finding common ground. *Archives of Internal Medicine*, 171(1), 56-65.
- Owolabi, T., & Simpson, I. (2012). Documenting and coding preventive visits: A physician's perspective. *Family Practice Management*. Retrieved May 6, 2017, from <http://www.aafp.org/fpm/2012/0700/p12.pdf>
- Patten, C. A., Choi, W. S., Vickers, K. S., & Pierce, J. P. (2001). Persistence of depressive symptoms in adolescents. *Neuropsychopharmacology*, 25, S89-S91.
- Richardson, L. P., Ludman, E., McCauley, E., Lindenbaum, J., Larson, C., Zhou, C., . . . Katon, W. (2014). Collaborative care for adolescents with depression in primary care: A randomized clinical trial. *Journal of the American Medical Association*, 312(8), 809-816.
- Richardson, L. P., McCauley, E., Grossman, D. C., McCarty, C. A., Richards, J., Russo, J. E., . . . Katon, W. (2010). Evaluation of the Patient Health Questionnaire-9 Item for detecting major depression among adolescents. *Pediatrics*, 126(6), 1117-1123.
- Rodríguez, G. (2013). Lecture notes on generalized linear models *Models for count data with overdispersion*: Princeton University.
- Rodríguez, G., & Elo, I. (2003). Intra-class correlation in random-effects models for binary data. *The Stata Journal*, 3(1), 32-46.
- Romano-Clarke, G., Tang, M. H., Xerras, D. C., Egan, H. S., Pasinski, R. C., Kamin, H. S., . . . Murphy, J. M. (2014). Have rates of behavioral health assessment and treatment increased for Massachusetts children since the Rosie D. decision? A report from two primary care practices. *Clinical Pediatrics*, 53(3), 243-249.
- Rushton, J., Bruckman, D., & Kelleher, K. (2002). Primary care referral of children with psychosocial problems. *Archives of Pediatrics and Adolescent Medicine*, 156(6), 592-598.

- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420.
- Simon, G. E., Rutter, C. M., Peterson, D., Oliver, M., Whiteside, U., Operskalski, B., & Ludman, E. J. (2013). Does response on the PHQ-9 depression questionnaire predict subsequent suicide attempt or suicide death? *Psychiatric Services*, *64*(12), 1195-1202.
- Siu, A. L. on behalf of the US Preventive Services Task Force (2016). Screening for depression in children and adolescents: US Preventive Services Task Force recommendation statement. *Pediatrics*, *137*(3), e20154467.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Patient Health Questionnaire Primary Care Study Group. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *Journal of the American Medical Association*, *282*(18), 1737-1744.
- Stevens, J., Wang, W., Fan, L., Edwards, M. C., Campo, J. V., & Gardner, W. (2009). Parental attitudes toward children's use of antidepressants and psychotherapy. *Journal of Child and Adolescent Psychopharmacology*, *19*(3), 289-296.
- Stice, E., Shaw, H., Bohon, C., Marti, C. N., & Rohde, P. (2009). A meta-analytic review of depression prevention programs for children and adolescents: Factors that predict magnitude of intervention effects. *Journal of Consulting and Clinical Psychology*, *77*(3), 486-503.
- Sudhanthar, S., Thakur, K., Sigal, Y., & Turner, J. (2015). Improving validated depression screen among adolescent population in primary care practice using electronic health records (EHR). *BMJ Quality Improvement Reports*, *4*(1), u209517-w203913.
- Suicide Prevention Resource Center. (2014). Tips and strategies for billing mental health services in a primary care setting. Retrieved May 28, 2018, from <https://www.sprc.org/sites/default/files/tipsandstrategiesforbilling.pdf>
- TeenScreen. (2011). TeenScreen primary care: Screening questionnaire overview: PSC-Y, PHQ-9, CRAFFT: TeenScreen Primary Care & TeenScreen National Center for Mental Health Checkups at Columbia University.
- Thomas, C. R., & Holzer, C. E. (2006). The continuing shortage of child and adolescent psychiatrists. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(9), 1023-1031.
- USPSTF. (2009). Screening and treatment for major depressive disorder in children and adolescents: US Preventive Services Task Force recommendation statement. *Pediatrics*, *123*(4), 1223-1228.
- USPSTF. (2016). Final recommendation statement: Depression in children and adolescents: Screening. Retrieved April 22, 2017, from https://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/depression-in-children-and-adolescents-screening_1
- Wang, P. S., Lane, M., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Twelve-month use of mental health services in the United States: Results from the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 629-640.
- Young, J. F., Bush, M. L., & McCarthy, A. E. (2018). Implications of AAP's new guidelines for adolescent depression: Part one. Retrieved from

- <https://policylab.chop.edu/blog/implications-aaps-new-guidelines-adolescent-depression-part-one>
- Young, J. F., Mufson, L., & Gallop, R. (2010). Preventing depression: A randomized trial of Interpersonal Psychotherapy-Adolescent Skills Training. *Depression & Anxiety, 27*(5), 426-433.
- Zenlea, I. S., Milliren, C. E., Mednick, L., & Rhodes, E. T. (2014). Depression screening in adolescents in the United States: A national study of ambulatory, office-based practice. *Academic Pediatrics, 14*(2), 186-191.
- Zito, J. M., Safer, D. J., Gardner, J. F., Magder, L., Soeken, K., Boles, M., . . . Riddle, M. A. (2003). Psychotropic practice patterns for youth: A 10-year perspective. *Archives of Pediatrics and Adolescent Medicine, 157*(1), 17-25.
- Zito, J. M., Safer, D. J., Zuckerman, I. H., Gardner, J. F., & Soeken, K. (2005). Effect of Medicaid eligibility category on racial disparities in the use of psychotropic medications among youths. *Psychiatric Services, 56*(2), 157-163.
- Ziv, A., Boulet, J. R., & Slap, G. B. (1999). Utilization of physician offices by adolescents in the United States. *Pediatrics, 104*(1), 35-42.
- Zuckerbrot, R. A., Cheung, A. H., Jensen, P. S., Stein, R. E., Laraque, D., & Glad-PC Steering Group. (2007). Guidelines for Adolescent Depression in Primary Care (GLAD-PC): I. Identification, assessment, and initial management. *Pediatrics, 120*(5), e1299-e1312.
- Zuckerbrot, R. A., Cheung, A. H., Jensen, P. S., Stein, R. E. K., Laraque, D., & GLAD-PC Steering Group. (2018). Guidelines for Adolescent Depression in Primary Care (GLAD-PC): Part I. Practice preparation, identification, assessment, and initial management. *Pediatrics, 141*(3), e20174081.
- Zuckerbrot, R. A., Maxon, L., Davies, M., Fisher, P. W., & Shaffer, D. (2007). Adolescent depression screening in primary care: Feasibility and acceptability. *Pediatrics, 119*(1), 101-108.

Table 1

Patient demographic characteristics

Demographic Characteristics	Full Sample						Selected Subsamples					
	Overall		Subthreshold Score (5-10)		Threshold Score (11-27)		Overall		Subthreshold Score (5-10)		Threshold Score (11-27)	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Sex												
Male	5,390	50.31	551	41.40	122	26.18	72	36.00	37	37.00	35	35.00
Female	5,323	49.69	780	58.60	344	73.82	128	64.00	63	63.00	65	65.00
Race												
American Indian or Alaska Native	7	.07	0	.00	0	.00	0	.00	0	.00	0	.00
Asian	293	2.73	39	2.93	13	2.79	8	4.00	6	6.00	2	2.00
Black or African American	2,969	27.71	448	33.66	181	38.84	76	38.00	41	41.00	35	35.00
Native Hawaiian or Other Pacific Islander	0	.00	0	.00	0	.00	0	.00	0	.00	0	.00
White	6,291	58.72	682	51.24	222	47.64	96	48.00	43	43.00	53	53.00
Multiple Races	70	.65	13	.98	3	.64	1	.50	1	1.00	0	.00
Other	1,079	10.07	149	11.19	47	10.09	19	9.50	9	9.00	10	10.00
Refused or Unknown	4	.0004	0	.00	0	.00	0	.00	0	.00	0	.00
Ethnicity												
Hispanic or Latino	547	5.11	90	6.76	38	8.15	10	5.00	2	2.00	8	8.00
Not Hispanic or Latino	10,141	10.07	1,239	93.09	427	91.63	190	95.00	98	98.00	92	92.00
Refused	25	.23	2	.15	1	.21	0	.00	0	.00	0	.00
Insurance type												

Private	8,089	75.50	908	68.22	279	59.87	129	64.50	65	65.00	64	64.00
Medicaid	2,496	23.30	405	30.43	183	39.27	68	34.00	32	32.00	36	36.00
Self-pay	128	1.19	18	1.35	4	.86	3	1.50	3	3.00	0	.00
Practice location												
Urban	2,943	27.47	485	36.44	175	37.55	75	37.50	37	37.00	38	38.00
Suburban	7,770	72.53	846	63.56	291	62.45	125	62.50	63	63.00	62	62.00

Table 2

Rates of screening at CHOP primary care practices

Practice Name	Patients Administered PHQ-9-M		
	N	%	95% CI
Practice 1	278	85.28	81.41, 89.14
Practice 2	236	81.94	77.48, 86.41
Practice 3	106	74.65	67.41, 81.89
Practice 4	545	87.06	84.42, 89.70
Practice 5	286	55.97	51.65, 60.29
Practice 6	231	91.67	88.23, 95.10
Practice 7	419	75.91	72.33, 79.48
Practice 8	238	81.79	77.33, 86.25
Practice 9	215	82.06	77.38, 86.74
Practice 10	175	71.14	65.44, 76.84
Practice 11	80	76.92	68.69, 85.16
Practice 12	465	80.03	76.77, 83.29
Practice 13	10	2.75	1.06, 4.43
Practice 14	340	91.89	89.10, 94.69
Practice 15	611	64.11	61.06, 67.16
Practice 16	247	88.85	85.13, 92.57
Practice 17	175	76.75	71.23, 82.28
Practice 18	258	62.77	58.08, 67.47
Practice 19	278	83.23	79.21, 87.26
Practice 20	59	64.84	54.84, 74.83
Practice 21	204	84.30	79.68, 88.91
Practice 22	0	0.00	N/A
Practice 23	155	79.90	74.21, 85.59
Practice 24	189	76.21	70.87, 81.55
Practice 25	20	4.64	2.65, 6.63
Practice 26	94	73.44	65.68, 81.19
Practice 27	237	68.30	63.38, 73.22
Practice 28	237	67.14	62.22, 72.06
Practice 29	5	1.26	.16, 2.37
Practice 30	305	71.26	66.96, 75.57
Practice 31	318	86.18	82.64, 89.72
<i>Total for all practices</i>	<i>7,016</i>	<i>65.49</i>	<i>64.59, 66.39</i>
<i>Total for practices with routine screening^a</i>	<i>6,981</i>	<i>76.25</i>	<i>75.38, 77.12</i>

^a Not including patients from Practices 13, 22, 25, and 29, where the electronic screening system was initiated later or an alternative behavioral health measure was being used during the study period.

Table 3

Rates of immediate PCP responses for patients with subthreshold and threshold PHQ-9-M scores

Response	Overall			Subthreshold Score (5-10)			Threshold Score (11-27)		
	N	%	95% CI	N	%	95% CI	N	%	95% CI
“Suicide risk high – Followed emergency procedures”	15	.83	.41, 1.26	2	.15	-.06, .36	13	2.79	1.29, 4.29
“Primary management of depression in primary care office”	206	11.46	9.99, 12.94	160	12.02	10.27, 13.77	46	9.87	7.15, 12.59
“Referred to behavioral health provider”	449	24.99	22.98, 26.99	289	21.71	19.50, 23.93	160	34.33	30.01, 38.66
“Did not refer – Already receiving behavioral health services”	326	18.14	16.36, 19.92	213	16.00	14.03, 17.98	113	24.25	20.34, 28.15
“No action needed – Reviewed questionnaire, inaccurate responses”	475	26.43	24.39, 28.47	444	33.36	30.82, 35.89	31	6.65	4.38, 8.92
Missing drop-down menu response	326	18.14	16.36, 19.92	223	16.75	14.75, 18.76	103	22.10	18.32, 25.88

Table 4

Follow-up over one year for patients with subthreshold and threshold PHQ-9-M scores according to EHR data extraction

	Overall			Subthreshold Score (5-10)			Threshold Score (11-27)		
	<i>N</i>	%	95% CI	<i>N</i>	%	95% CI	<i>N</i>	%	95% CI
Documented follow-up type									
Behavioral health referral									
Any referral	546	30.38	28.26, 32.51	344	25.85	23.49, 28.20	202	43.35	38.83, 47.86
Drop-down menu referral	439	24.99	22.98, 26.99	289	21.71	19.50, 23.93	160	34.33	30.01, 38.66
Other documentation of referral	97	5.40	4.35, 6.44	55	4.13	3.06, 5.20	42	9.01	6.40, 11.62
Depressive disorder diagnosis	425	23.65	21.68, 25.62	211	15.85	13.89, 17.82	214	45.92	41.38, 50.46
Antidepressant medication	254	14.13	12.52, 15.75	132	9.92	8.31, 11.53	122	26.18	22.17, 30.19
PHQ-9-M re-administration	52	2.89	2.12, 3.67	36	2.70	1.83, 3.58	16	3.43	1.77, 5.09
None of the above	935	52.03	49.72, 54.34	818	61.46	58.84, 64.08	117	25.11	21.16, 29.06
Number of follow-up types									
None	913	50.81	48.49, 53.12	799	60.03	57.40, 62.67	114	24.46	20.55, 28.28
One type	545	30.33	28.20, 32.46	365	27.42	25.02, 29.82	180	38.63	34.19, 43.06

Two types	290	16.14	14.44, 17.84	145	10.89	9.22, 12.57	145	31.12	26.90, 35.33
Three types	44	2.45	1.73, 3.16	20	1.50	.85, 2.16	24	5.15	3.14, 7.16
Four types	5	.28	.03, .52	2	.15	.06, .36	3	.64	.09, 1.37

Table 5

Follow-up over one year for selected subsamples of patients with subthreshold and threshold PHQ-9-M scores according to manual EHR chart review

	Overall			Subthreshold Score (5-10)			Threshold Score (11-27)		
	<i>N</i>	%	95% CI	<i>N</i>	%	95% CI	<i>N</i>	%	95% CI
Behavioral health referral									
Any referral	111	55.50	48.55, 62.45	37	37.00	27.37, 46.63	74	74.00	65.25, 82.75
Drop-down menu referral	69	34.50	27.85, 41.15	23	23.00	14.61, 31.39	46	46.00	36.06, 55.94
Other documentation of referral	101	50.50	43.51, 57.49	33	33.00	23.62, 42.38	68	68.00	58.70, 77.30
Both types of referral	59	29.50	23.13, 35.87	19	19.00	11.18, 26.82	40	40.00	30.23, 49.77
Assessment and diagnosis									
Diagnostic assessment	3	1.50	-.20, 3.20	1	1.00	-.98, 2.98	2	2.00	-.79, 4.79
Depressive disorder diagnosis	54	27.00	20.79, 33.21	16	16.00	8.69, 23.31	38	38.00	28.32, 47.68
Supportive monitoring and care coordination									
Social work encounter on date of index well visit	13	6.50	3.05, 9.95	2	2.00	-.79, 4.79	11	11.00	4.76, 17.24

Psychoeducation	44	22.00	16.21, 27.79	16	16.00	8.69, 23.31	28	28.00	19.05, 36.95
Evidence-based treatment recommendation	2	1.00	-.39, 2.39	1	1.00	-.98, 2.98	1	1.00	-.98, 2.98
Treatment plan	9	4.50	1.60, 7.40	1	1.00	-.98, 2.98	8	8.00	2.59, 13.41
Safety plan	23	11.50	7.04, 15.96	5	5.00	.65, 9.35	18	18.00	10.34, 25.66
Support and monitoring	70	35.00	28.33, 41.67	23	23.00	14.61, 31.39	47	47.00	37.05, 56.95
Consultation with mental health provider	8	4.00	1.26, 6.74	2	2.00	-.79, 4.79	6	6.00	1.26, 10.74
Follow-up with mental health provider	17	8.50	4.60, 12.40	3	3.00	-.40, 6.40	14	14.00	7.08, 20.92
PHQ-9-M Re- administration	4	2.00	.04, 3.96	3	3.00	-.40, 6.40	1	1.00	-.98, 2.98
Administration of other depression measure	21	10.50	6.21, 14.79	10	10.00	4.02, 15.98	11	11.00	4.76, 17.24
Treatment									
Psychotherapy	65	32.50	25.95, 39.05	22	22.00	13.74, 30.26	43	43.00	33.13, 52.87
Antidepressant medication	47	23.50	17.57, 29.43	14	14.00	7.08, 20.92	33	33.00	23.62, 42.38
Antidepressant medication monitoring	25	12.50	7.88, 17.12	7	7.00	1.91, 12.09	18	18.00	10.34, 26.55
Emergency department	17	8.50	4.60, 12.40	3	3.00	-.40, 6.40	14	14.00	7.08, 20.92

Inpatient hospitalization	12	6.00	2.68, 9.32	3	3.00	-.40, 6.40	9	9.00	3.29, 14.71
None of the above	46	23.00	17.12, 28.88	37	37.00	27.37, 46.63	9	9.00	3.29, 14.71

Table 6

Patient demographic characteristics as predictors of PHQ-9-M score with square root transformation

Predictor Variable	<i>b</i> (SE)	β	<i>t</i>-value	<i>p</i>-value	95% CI for <i>b</i>
Constant	.99 (.02)		43.81	<.001	.95, 1.04
Female vs. male	.39 (.03)	.17	14.18	<.001	.33, .44
Minority vs. majority racial/ethnic group	.14 (.03)	.06	4.39	<.001	.08, .21
Medicaid vs. other insurance	.34 (.04)	.12	9.73	<.001	.27, .41
Urban vs. suburban	.15 (.04)	.06	4.29	<.001	.08, .22

Note. $R^2=.06$, Adjusted $R^2=.06$, $F(4, 6907)=118.28$, $p<.001$.

Table 7

PHQ-9-M score, suicide risk, and patient demographic characteristics as predictors of immediate PCP response

Response ^a	<i>b</i> (<i>SE</i>)	Wald χ^2	<i>p</i> -value	95% CI for Odds Ratios		
				Lower	Odds Ratio	Upper
“Suicide risk high – Followed emergency procedures”						
Intercept	-27.67 (1.32)	440.40	<.001			
PHQ-9-M score	.35 (.06)	31.88	<.001	1.25	1.42	1.60
Suicide risk	3.75 (1.09)	11.81	.001	5.00	42.35	358.54
Female vs. male ^b	--	--	--	--	--	--
Minority vs. majority racial/ethnic group	-.02 (.72)	.001	.98	.24	.98	3.98
Medicaid vs. other insurance	-.32 (.62)	.28	.60	.22	.72	2.42
Urban vs. suburban practice	1.29 (.72)	3.19	.07	.88	3.62	14.88
“Primary management of depression in primary care office”						
Intercept	-2.50 (.28)	81.49	<.001			
PHQ-9-M score	.17 (.03)	32.05	<.001	1.12	1.19	1.26
Suicide risk	.73 (.27)	7.59	.01	1.23	2.07	3.48
Female vs. male	.03 (.17)	.03	.85	.74	1.03	1.45
Minority vs. majority racial/ethnic group	.26 (.20)	1.67	.20	.87	1.30	1.94

^a The reference category was “No action needed – Reviewed questionnaire, inaccurate responses.”

^b Sex was not applicable as a predictor of “Suicide risk high – Followed emergency procedures” versus “No action needed – Reviewed questionnaire, inaccurate responses” because only female patients required emergency procedures due to high suicide risk.

Medicaid vs. other insurance	-.03 (.20)	.02	.88	.66	.97	1.43
Urban vs. suburban practice	.24 (.20)	1.42	.23	.86	1.27	1.89
“Referred to behavioral health provider”						
Intercept	-2.35 (.24)	97.01	<.001			
PHQ-9-M score	.19 (.03)	51.50	<.001	1.15	1.21	1.27
Suicide risk	1.52 (.22)	49.89	<.001	3.00	4.57	6.97
Female vs. male	.49 (.15)	11.22	.001	1.23	1.63	2.18
Minority vs. majority racial/ethnic group	.20 (.17)	1.36	.24	.88	1.22	1.69
Medicaid vs. other insurance	.36 (.16)	4.99	.03	1.05	1.43	1.97
Urban vs. suburban practice	-.08 (.17)	.24	.63	.66	.92	1.29
“Did not refer – Already receiving behavioral health services”						
Intercept	-2.13 (.25)	75.21	<.001			
PHQ-9-M score	.19 (.03)	47.33	<.001	1.15	1.21	1.28
Suicide risk	1.66 (.23)	54.27	<.001	3.39	5.28	8.21
Female vs. male	.38 (.16)	5.55	.02	1.07	1.46	2.00
Minority vs. majority racial/ethnic group	-.57 (.18)	9.49	.002	.40	.57	.81
Medicaid vs. other insurance	.28 (.18)	2.26	.13	.92	1.32	1.88
Urban vs. suburban practice	-.51 (.20)	6.42	.01	.41	.60	.89

Note. Cox and Snell R^2 =.24; Nagelkerke R^2 =.25, $\chi^2(24)$ =393.27, p <.001.

Table 8

PHQ-9-M score, suicide risk, and patient demographic characteristics as predictors of one-year follow-up according to EHR data extraction

Response	<i>b</i> (<i>SE</i>)	Wald χ^2	<i>p</i> -value	95% CI for Odds Ratios		
				Lower	Odds Ratio	Upper
Depressive disorder diagnosis						
Constant	-2.70 (.18)	232.29	<.001		.07	
PHQ-9-M score	.13 (.02)	62.14	<.001	1.10	1.14	1.17
Suicide risk	1.17 (.14)	74.35	<.001	2.47	3.22	4.21
Female vs. male	.21 (.13)	2.55	.11	.95	1.23	1.60
Minority vs. majority racial/ethnic group	-.08 (.14)	.34	.56	.70	.92	1.22
Medicaid vs. other insurance	-.09 (.14)	.39	.53	.70	.92	1.21
Urban vs. suburban practice	-.44 (.15)	8.55	.003	.48	.64	.86
Behavioral health referral						
Constant	-1.97 (.16)	154.99	<.001		.14	
PHQ-9-M score	.06 (.02)	15.92	<.001	1.03	1.06	1.09
Suicide risk	.53 (.13)	17.14	<.001	1.32	1.70	2.18
Female vs. male	.28 (.11)	6.06	.01	1.06	1.33	1.66
Minority vs. majority racial/ethnic group	.30 (.13)	5.49	.02	1.05	1.34	1.72
Medicaid vs. other insurance	.17 (.12)	1.99	.16	.94	1.18	1.50
Urban vs. suburban practice	-.04 (.13)	.10	.75	.75	.96	1.24
Antidepressant medication						

Constant	-2.78 (.20)	191.34	<.001		.06	
PHQ-9-M score	.11 (.02)	38.01	<.001	1.08	1.12	1.16
Suicide risk	.71 (.17)	17.73	<.001	1.46	2.03	2.82
Female vs. male	.35 (.16)	4.81	.03	1.04	1.42	1.95
Minority vs. majority racial/ethnic group	-.92 (.18)	27.20	<.001	.28	.40	.56
Medicaid vs. other insurance	-.19 (.18)	1.16	.28	.59	.83	1.17
Urban vs. suburban practice	-.53 (.20)	7.25	.01	.40	.59	.87
PHQ-9-M re-administration						
Constant	-4.16 (.43)	94.88	<.001		.02	
PHQ-9-M score	.04 (.04)	.89	.35	.96	1.04	1.12
Suicide risk	-.16 (.36)	.21	.64	.42	.85	1.70
Female vs. male	.50 (.32)	2.46	.12	.88	1.66	3.11
Minority vs. majority racial/ethnic group	-.74 (.34)	4.63	.03	.25	.48	.94
Medicaid vs. other insurance	-.72 (.35)	4.11	.04	.24	.49	.98
Urban vs. suburban practice	1.31 (.34)	14.67	<.001	1.89	3.69	7.19

Note. Depressive disorder diagnosis model: Cox and Snell $R^2=.14$, Nagelkerke $R^2=.22$, $\chi^2(6)=278.88$, $p<.001$. Behavioral health referral model: Cox and Snell $R^2=.05$, Nagelkerke $R^2=.07$, $\chi^2(6)=88.69$, $p<.001$. Antidepressant medication prescription model: Cox and Snell $R^2=.10$, Nagelkerke $R^2=.18$, $\chi^2(6)=184.50$, $p<.001$. PHQ-9-M re-administration model: Cox and Snell $R^2=.01$, Nagelkerke $R^2=.05$, $\chi^2(3)=20.72$, $p=.002$.

Table 9

PHQ-9-M score, suicide risk, and patient demographic characteristics as predictors of total number of follow-up types according to manual EHR chart review

Predictor Variable	<i>b</i> (<i>SE</i>)	Wald χ^2	<i>p</i> -value	95% CI for Odds Ratios		
				Lower	Odds Ratio	Upper
Intercept	.13 (.21)	.37	.54	.75	1.14	1.73
PHQ-9-M score	.06 (.02)	14.44	<.001	1.03	1.07	1.10
Suicide risk	.67 (.15)	19.66	<.001	1.45	1.96	2.63
Female vs. male	.07 (.14)	.22	.64	.81	1.07	1.42
Minority vs. majority racial/ethnic group	-.25 (.15)	2.49	.12	.58	.78	1.06
Medicaid vs. other insurance	.18 (.15)	1.49	.22	.90	1.20	1.60
Urban vs. suburban	.10 (.16)	.39	.53	.81	1.10	1.51

Note. Pseudo R^2 =.05 (Heinzl & Mittlböck, 2003). Model $\chi^2(6)$ =62.63, p <.001.

Appendix

Data on Adolescent Patients Attending Non-Age 16 Well Visits

Table 1

Rates of completed PHQ-9-M screens and elevated depressive symptoms by age group

Age Group	Screening Rate			Screened Patients with Subthreshold Score (5-10)			Screened Patients with Threshold Score (11-27)		
	<i>N</i>	%	95% CI	<i>N</i>	%	95% CI	<i>N</i>	%	95% CI
12.00-12.99 years old	496	2.54	2.32, 2.77	75	15.12	11.96, 18.28	21	4.23	2.46, 6.01
13.00-13.99 years old	485	4.07	3.71, 4.42	82	16.91	13.56, 20.25	21	4.33	2.51, 6.15
14.00-14.99 years old	2,576	23.38	22.59, 24.17	431	16.73	15.29, 18.17	141	5.47	4.59, 6.35
15.00-15.49 years old	1,930	35.80	34.52, 37.08	333	17.25	15.57, 18.94	108	5.59	4.57, 6.62
15.50-16.49 years old (full sample)	6,923	64.62	63.72, 65.53	1,331	19.23	18.30, 20.15	466	6.73	6.14, 7.32
16.50-16.99 years old	142	3.39	2.84, 3.93	30	21.13	14.33, 27.92	9	6.34	2.28, 10.39
17.00-17.99 years old	138	1.72	1.43, 2.003	22	15.94	9.76, 22.13	11	7.97	3.40, 12.55
<i>Total</i>	12,690	17.93	17.65, 18.22	2,304	18.16	17.49, 18.83	777	6.12	5.71, 6.54