Screening for depression with the PHQ-9 in young adults affected by HIV

Laura Mufson, PhD[a,*], **Corey Morrison**[b], **Eileen Shea, MPH**[c], **Luke Kluisza, MS**[b], **Reuben Robbins, PhD**[b], **Ying Chen, MD**[d,e], **Claude A. Mellins, PhD**[b]

[a]Child and Adolescent Psychiatry, New York State Psychiatric Institute and Columbia University, 1051 Riverside Drive, New York, NY 10032, USA

[b]HIV Center for Clinical and Behavioral Studies, New York State Psychiatric Institute and Columbia University, 1051 Riverside Dr #15, New York, NY 10032, USA

[c]Department of Psychiatry, Mental Health Data Science, Columbia University Medical Center, 1051 Riverside Dr, New York, NY 10032, USA

[d]Mailman School of Public Health, Columbia University, 722 W 168th St, New York, NY 10032

[e]New York State Psychiatric Institute and Columbia University, 1051 Riverside Dr #15, New York, NY 10032, USA

Abstract

Background—Major depressive disorder (MDD) is common among young adults with perinatally acquired HIV (YAPHIV), however it is often underdiagnosed and untreated. The PHQ-9 and PHQ-2 are widely used screening instruments for MDD. This study evaluates the accuracy of recommended PHQ-9 and PHQ-2 cut-scores of 10 and 3 for YAPHIV and YA who were perinatally HIV exposed but uninfected (YAPHEU).

Methods—The PHQ-9 was administered to participants (n=203) in a longitudinal cohort study using the DISC-IV as the gold standard for diagnosing depression. PHQ-9 and PHQ-2 sensitivity and specificity were calculated. ROC curves were constructed for the overall sample and YAPHIV and YAPHEU subsamples.

Results—Almost all participants were Black and Latinx, ages 18–29. Overall, the recommended PHQ-9 cut-score of 10 yielded a sensitivity of .47 (95%CI [0.23,0.72]) and specificity of .86 (95%CI [0.80, 0.91]). Results indicate that PHQ-9 cut-scores of 7 and PHQ-2 cut-scores of 2 increased sensitivity to .76 (95%CI [0.50, 0.93]) and .71(95%CI [0.44,0.90]), and decreased specificity to .72 (95%CI [0.65, 0.79]) and .73 (95%CI [0.66, 0.79]) respectively. Among subsamples, existing PHQ-9 cut-scores were more accurate for MDD diagnoses in YAPHEU (N=11) than YAPHIV(N=6). No race/ethnicity or age differences were found.

Limitations—Participants were recruited from clinics in NYC and may not reflect all YAPHIV and YAPHEU. Without a white HIV comparison group, no conclusions could be made on the impact of race/ethnicity on optimized PHQ-9 cut-scores.

[*]**Corresponding Author.** laura.mufson@nyspi.columbia.edu. Tel/Fax: (646) 774-5791, Address: Child and Adolescent Psychiatry, New York State Psychiatric Institute and Columbia University, 1051 Riverside Drive, New York, NY 10032, USA.

Conclusions—Using tailored cut scores for HIV-affected populations may increase identification of those experiencing or at risk for MDD. Given the need for increased depression screening in HIV care, use of optimized cut-scores could benefit at-risk populations in the US and globally.

Introduction

Depression is a leading cause of disability worldwide and, as of 2017, is a leading contributor to the global burden of disease (James et al., 2018). Mental health problems, particularly depression, are recognized as a significant barrier to achieving success in the effort to end the HIV epidemic (Remien et al., 2019). Numerous studies have shown increased risk for depression in people living with or at risk for HIV (PLHIV), including adolescents and young adults with perinatally acquired HIV (Bhatia and Munjal, 2014; Do et al., 2014; Mellins and Malee, 2013). Major depressive disorder (MDD) is also a significant predictor of poor health outcomes in PLHIV (Bing et al., 2001; Lopes et al., 2011; Owe-Larsson et al., 2009; Simoni et al., 2011). New York City (NYC) is home to approximately 20% of the 12,310 cases of young people with perinatally acquired HIV (PHIV) in the US (Centers for Disease Control and Prevention, 2020), the vast majority of whom are entering or are well into young adulthood. The developmental period of young adulthood is increasingly recognized as a particularly vulnerable time for the onset of depression (Kessler et al., 2005). Thus, identification of depression in young adults with PHIV (YAPHIV) could have a significant impact on their quality of life. The majority of YAPHIV in the US are Black and Latinx, from vulnerable communities affected by poverty, neighborhood violence, substance abuse and health care disparities, placing them at a greater risk for depression and worse health outcomes (Centers for Disease Control and Prevention, 2020; Kang et al., 2011; Kang et al., 2021). Yet, in spite of numerous recommendations for mental health screening and treatment programs for young people living with HIV in the US and globally (Mellins and Malee, 2013; Remien et al., 2019), identification and treatment are still challenging, as few evidence-based intervention programs exist for this population (Bhana et al., in press)

Depression among PLHIV is often unrecognized and untreated because of difficulties accessing assessment and treatment due to overextended medical clinics, lack of prioritization of mental health treatment, and often few resources for screening and evaluation (Bhana et al., in press; Remien et al., 2019). These problems parallel those in the general community. Several U.S. studies estimated that only 1–2% of the United States' general primary care population has received depression screening (Harrison et al., 2010; Margaretten et al., 2013; McGoey et al., 2013). Depression is often underdiagnosed and underrecognized specifically among Black and Latinx populations (Borowsky et al., 2000; Skaer et al., 2000; Stockdale et al., 2008), exacerbating racial disparities in mental health care. Clinical presentations of depression symptoms have been found to differ among racial and ethnic groups (Brown et al., 1996; Das et al., 2006; Kirmayer, 2001), as well as groups with chronic health conditions (Freedland et al., 1992; Kalichman et al., 2000), further underscoring the need for well-validated screening instruments with population-specific cut-scores. While many depression screening tools exist, few have been

developed with population-specific cut-scores culturally relevant for the HIV and Black and Latinx communities (Azocar et al., 2001; Huang et al., 2006; Posner et al., 2001).

In response to recommendations for depression screening (Calonge et al., 2009), the Patient Health Questionnaire (PHQ-9) has increasingly become one of the most widely used brief, self-report screening tools for identifying those at risk for MDD in the general population (Kroenke and Spitzer, 2002), as well as in studies of PLHIV (Monahan et al., 2009). The scale comprises nine items reflecting the Diagnostic and Statistical Manual of Mental Disorders (DSM) (American Psychiatric Association, 2013) defined symptoms for MDD. A cut-point of 10 is typically used to delineate symptoms indicative of an MDD diagnosis and has been validated in a number of studies (Kroenke et al., 2001). Prior validation studies have shown the PHQ-9 to have a sensitivity of 88% and a specificity of 88% for MDD (Kroenke et al., 2001). The PHQ-2, consisting of the first two items of the PHQ-9 with the same rating scale, focuses on frequency of depressed mood and anhedonia, and is often administered as a quicker initial step in depression screening to identify individuals who require further evaluation for depressive symptoms (Löwe et al., 2005). A PHQ-2 score ≥ 3 has 83% sensitivity and 92% specificity for MDD (Kroenke et al., 2003). Clinical cut-scores (≥ 10 on the PHQ-9 and ≥ 3 on the PHQ-2) are considered reflective of risk for MDD according to studies with largely White, middle class individuals (Manea et al., 2012). Unfortunately, it is not clear whether these clinical cut-scores are appropriate for those most in need of clinical care in populations living with or affected by HIV, most of whom are from Black and Latinx backgrounds (Centers for Disease Control and Prevention, 2020), including those who are young adults.

To our knowledge there are no studies examining the accuracy of recommended PHQ-9 and PHQ-2 cut-scores in people living with HIV, including YAPHIV, in spite of the recommendations for routine depression screening for patients living with HIV (Collins et al., 2013; Remien et al., 2019; Tomlinson et al., 2009). In the US and across the globe, there is a lack of knowledge about the clinically relevant cut-scores to identify those in need of care in diverse communities. Tools that can be administered by lay staff, with easy to define clinical cut-scores will be important for facilitating and disseminating screening as a routine part of health care in resource constrained settings.

To address this knowledge gap, we examined the accuracy of the PHQ-9 cut-score of 10 and the PHQ-2 cut-score of 3 to detect DSM-IV defined MDD in a sample of YAPHIV and a comparison group of young adults perinatally HIV exposed but uninfected (YAPHEU) from similar communities, in a longitudinal cohort study on the psychiatric and social impact of perinatally acquired HIV infection (Mellins et al., 2009; Nguyen et al., 2020). Specifically, this paper examines the sensitivity and specificity of PHQ-9 and PHQ-2 cut-scores in relation to MDD diagnosis on the Diagnostic Interview Schedule for Children (DISC-IV) version for young adults (Shaffer et al., 2000). The DISC is a well-validated structured interview for the assessment of psychiatric disorders.

Methods

The Child and Adolescent Self-Awareness and Health study (CASAH) is an ongoing longitudinal study of a cohort (N=340) of young people living with PHIV (n=206) and PHEU, who were originally recruited when they were 9–16 years old from four NYC medical centers between 2003–2008. The primary aims of CASAH are to examine psychiatric and psychosocial outcomes in this population and their association with HIV, young adult transition milestones, and other social determinants of health. Details on recruitment have been previously presented (Mellins et al., 2009; Mellins et al., 2012). In brief, we approached 83% of eligible participants and were able to recruit and enroll 93% of them (77% of all eligible participants). Inclusion criteria at baseline were perinatal HIV exposure; cognitive capacity to complete the interview; English-or Spanish-speaking; and a caregiver with legal capacity to sign consent for youth participation. Providers identified eligible patients in their clinics and referred interested caregivers and youth to the study team. Caregivers and youth provided consent (caregivers for youth <18 years, youth assent <18 years, and consent >18 years) and completed an assessment battery at enrollment and 18-months follow-up (FU). CASAH received additional funding to follow the cohort as they aged from adolescence into young adulthood and is currently in its 17th year. Participants continued to complete psychosocial batteries every 12–18 months administered by the research team, consisting of bachelor level lay staff from diverse racial and ethnic backgrounds and gender. The data for these analyses were taken from the seventh follow-up interview (FU7), when the participants were ages 20–29 and the PHQ-9 was first added to the battery. This study was approved by the Columbia University–New York State Psychiatric Institute Institutional Review Board. All young adults provided consent and were reimbursed for their time and travel.

Measures

The DISC-IV was used to assess MDD. The DISC-IV, which includes versions for adolescents and young adults, is a well-validated, structured instrument administered by highly trained research assistants to assess common psychiatric diagnoses (Shaffer et al., 2000) as defined by the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders system (DSM) (American Psychiatric Association, 2013). Youth were interviewed about their experience of depressive symptoms over the past year and met criteria for MDD as assessed by the DISC-IV young adult version. The DISC-IV has been used and validated across diverse populations, including Black and Latinx populations (Kim et al., 2003; Shaffer et al., 2000; Yeh et al., 2003).

At FU7, youth were also administered the PHQ-9 (Kroenke and Spitzer, 2002). The scale comprises nine items reflecting DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition) criteria for MDD (e.g., "little interest or pleasure in doing things"). Participants rate each item on a 4-point Likert scale ranging from 0 (not at all) to 3 (nearly every day). Total scores are then calculated, with higher scores indicating higher depression. The PHQ-2 is also scored on the same 4-point Likert scale as the PHQ-9, with higher total scores indicating higher depression.

Both the DISC-IV and the PHQ 9 were administered as part of an extensive interview battery. Neither the DISC-IV nor the PHQ 9 involve interviewer interpretation. The DISC and PHQ-9 are given in the same interview by the same person so there is no way to blind them from each other.

Analysis

Sensitivity and specificity of the PHQ-9 and PHQ-2 were first calculated based on established cut-scores for identification of MDD symptom levels (PHQ-9 score 10; PHQ-2 score 3), using DISC-IV diagnoses (diagnosis of MDD alone; diagnosis of MDD or dysthymia) as the gold standard. ROC (Receiver Operating Characteristic) curves were then constructed using DISC-IV MDD diagnosis as the binary classifier and either the PHQ-9 or the PHQ-2 as the diagnostic test of interest. ROC analysis involved calculating AUC (Area Under the Curve) as a measure of the accuracy of the test and determining the optimal cut-score for the diagnostic test using Youden's index (J statistic; Sensitivity + Specificity – 1). This analysis was completed in both the overall CASAH population, as well as in the YAPHIV and YAPHEU populations separately. To further understand differences between the YAPHIV and YAPHEU populations, differential rates of PHQ-9 item endorsement by HIV status were assessed using Chi-square and Fisher's exact tests.

Results

This analysis focused on 203 CASAH participants from FU7, who had DISC-IV information as well as a completed PHQ-9 (3 participants were excluded due to missing data). The population was also predominantly Black and/or Latinx (45% Black, 38% Latinx, 12% Black Hispanic, 5% Other), reflecting the full CASAH cohort and the epidemic in women and children in NYC at the time of recruitment (New York City Department of Health and Mental Hygiene, 2007). There were no significant differences between YAPHIV and YAPHEU population regarding ethnicity (p=0.67). Participants had an average age of 24.1 years (SD = 2.6) and were 54% female and 46% male. The mean age of YAPHIV was 24.4 (SD=2.6)) and significantly older than YAPHEU whose mean age was 23.6 (SD=2.6), p=.02. There was no significant difference for sex (p=0.94). YAPHIV represented nearly two-thirds of the sample (n=126, 62%), with YAPHEU comprising the remaining 38% (n=77). There were significantly more YAPHEU diagnosed with MDD (p=.02) and MDD + Dysthymia (p=.03) than YAPHIV.

Overall Sample:

Overall, the DISC conferred a diagnosis of MDD (N= 17) and a diagnosis of MDD+Dysthymia (N=18) in this sample. Using the existing cut-score of 10, we identified eight people who screened positive for MDD on the PHQ-9. Using a cut-score of 3, we identified eight people positive for MDD on the PHQ-3. With the new cut-score of 7, we identified 13 people as positive for MDD on the PHQ-9 and using the new cut-score of 2 on the PHQ-3, we identified 12 people as positive for MDD. In this combined population of YAPHIV and YAPHEU, using the existing cut-score of 10 or greater on the PHQ-9 as a screening tool for MDD resulted in a sensitivity of 0.47 and specificity of 0.86; the existing cut-score of 3 or greater on the PHQ-2 also resulted in these same sensitivity

and specificity values (Table 1). Sensitivity was minimally improved when considering a diagnosis of either MDD or dysthymia on the DISC as the reference standard. Furthermore, AUC values gathered from the ROC curves were only slightly larger when considering MDD and dysthymia together as a binary classifier of depression (PHQ-9: AUC = 0.80 95%CI [0.68;0.93]; PHQ-2: AUC = 0.76, 95%CI [0.62;0.89]) than when considering MDD alone (PHQ-9: AUC = 0.79, 95%CI [0.66,0.92]); PHQ-2: AUC = 0.47, 95%CI [0.60;0.88]) (Figure 1).

Optimal cut-score analysis using the ROC curves and Youden's index suggested that lowering the cut-scores for both the PHQ-9 and PHQ-2 below the existing cut-scores of 10 and 3, respectively, would improve sensitivity (Table 1, Figure 2). Using a PHQ-9 cut-score of 7 or greater as a measure of MDD increased sensitivity to 0.76 (95% CI [0.50,0.93]), while using a PHQ-2 cut-score of 2 or greater increased sensitivity to 0.71 (95%CI [0.44,0.90]). Since there is always a trade-off between sensitivity and specificity, specificity subsequently goes down for both the PHQ-9 and PHQ-2 with these lower cut-scores, although the optimal cut-score analysis considers the trade-off between increasing the true positive rate (TPR, i.e., sensitivity) while increasing the false positive rate (FPR, i.e., 1 − specificity).

Stratified Sample:

Stratification of the ROC analysis by HIV status shows differential accuracy of the PHQ questionnaires among YAPHIV and YAPHEU. Table 2 shows that among YAPHIV, the existing cut-scores are not capturing any of the individuals with an MDD diagnosis (N= 11)(i.e., sensitivity is zero). While lowering the cut-scores on the PHQ-9 and PHQ-2 to 7 and 2, respectively, improves sensitivity, particularly in the case of the PHQ-9, the AUC values (0.67(95%CI [0.58,0.75]) and 0.57 (95%CI [0.48,0.66]), respectively) indicate poor accuracy of the tests for MDD diagnosis among YAPHIV (N=6).

Comparatively, among YAPHEU (Table 3) we see good accuracy of both the PHQ-9 and PHQ-2 for MDD diagnosis with AUC values of 0.86 (95%CI [0.77,0.93]) and 0.84 (95%CI [0.73,0.91]), respectively. Among YAPHEU, the existing cut-scores on the PHQ-9 perform better for MDD diagnosis than they do among the pooled YAPHIV and YAPHEU populations (sensitivity = 0.73 (95%CI [0.39,0.94]), specificity = 0.88 (95%CI [0.78,0.95])); however, the optimal cut-score analysis still advocates for a lower threshold of a score of 7 for MDD screening in the YAPHEU population. For the PHQ-2, the existing cut-score and the optimal cut- score are the same in this population based on sensitivity and specificity.

Stratification of the ROC analysis by age category (20–24 vs. 25+), race (Black vs. Not Black), and ethnicity (Latinx vs. Not Latinx) also took place, with no meaningful differences detected.

Item Response Differences by HIV Status:

We looked at item response differences on the PHQ-9 by HIV status and found that the fourth question of the PHQ-9 was endorsed differently by YAPHIV and YAPHEU ($\chi^2 = 10.19$, df = 3, p = 0.02). This item asks about feeling tired or having little energy over the

last two weeks; YAPHIV on this item were more likely to respond, "Several days" (score of 1 on the Likert scale) compared to YAPHEU (45% vs. 30%, p = 0.03), while YAPHEU were more likely to respond, "Nearly every day" (score of 3 on the Likert scale) compared to YAPHIV (19% vs. 9%, p = 0.03). No other PHQ-9 items showed significant differences in response rates by HIV status.

Discussion

This study examines the diagnostic ability of recommended PHQ-9 and PHQ-2 cut-scores (10 and 3 respectively) to identify MDD among primarily Black and Latinx young adults who either grew up with PHIV or PHEU using a well validated psychiatric interview (DISC-IV) as the gold standard comparison. The findings suggest that among Black and Latinx young adults affected by HIV, the commonly recommended PHQ-9 and PHQ-2 cut-scores may miss a large proportion of individuals with MDD or dysthymia. Lowering the cut-scores to 7 (PHQ-9) and 2 (PHQ-2) substantially increased the tool's sensitivity to detect individuals with these disorders among our participants, although it also decreased specificity. For purposes of screening young adults to identify those who could benefit from treatment to prevent the negative sequelae of depression, higher sensitivity likely outweighs the detection of false positives. Additionally, the decrease in specificity from .86 to .73 in our study is not meaningful enough to be of concern if the goal of screening is to identify those at risk by including some with milder symptoms or transient distress who may not meet full criteria for MDD. Identifying these young adults early and linking them to services may be critical for prevention of more severe future episodes of depression and may also promote positive health outcomes. Numerous studies have shown an association between depression, non-adherence to antiretroviral treatment, and lack of viral suppression in adults and youth (Remien et al., 2019; Uthman et al., 2014; Wagner et al., 2020), further underscoring the need for depression screening in this population.

Interestingly, when we examined HIV status group differences, the standard cut-scores yielded better sensitivity and specificity among YAPHEU, in addition to the AUC values being higher among this group. This indicates that the PHQ questionnaires may perform better among certain populations, and that for groups like YAPHIV, even lower cut-scores may be necessary to reach a larger number of individuals at risk for severe illness. The difference in findings between the YAPHIV and the YAPHEU samples raises the question of how growing up with HIV infection impacts the reporting of depression symptoms. As found in our item analysis, YAPHEU were more likely to endorse more severe experience of fatigue than the YAPHIV subsample on the PHQ-9. Fatigue is one of the most commonly cited side-effects of antiretroviral treatment (Barroso et al., 2015). While this is inconsistent with prior research showing an inflation of somatic symptoms in those with HIV (Kalichman, Rompa, & Cage, 2000) and other chronic illnesses (Patterson et al., 2011), it is consistent with the conclusions that co-occurring somatic symptoms may complicate depression screening and diagnosis in populations with HIV and other chronic illnesses. Future research should also address this limitation in depression screening tools.

Limitations

The vast majority of study participants were Black and Latinx and given the lack of a white or other racial/ethnic HIV comparison groups, we are unable to draw conclusions on the impact of race and ethnicity on the optimized PHQ-9 cut-scores. We also do not have information on those we did not recruit which might inform whether there are any recruitment biases that have affected group differences or our study findings. Additionally, there was a relatively small number of participants who met criteria for a major depression diagnosis on the DISC-IV which weakened the power to detect the predictive validity of the cut-scores for identifying MDD.

Conclusion

The results suggest that recommended cut-scores on a screening measure such as the PHQ-9 might lack accuracy in certain populations. Therefore, it may be important to tailor cut-scores for the population being studied or served; using tailored cut-scores on the PHQ-9 and PHQ-2 optimized for a YAPHIV and YAPHEU sample would increase identification of young adults both at risk for depression and those currently experiencing MDD. Depression is often unrecognized and untreated for young adults living with HIV in the U.S., and it is an even greater public health crisis in low and middle-income countries where the vast majority of young people growing up with HIV or who are HIV-affected reside (UNAIDS, 2021). In sub-Saharan Africa, where 84% of youth living with HIV reside (United Nations Children's Fund (UNICEF), 2017), the reported prevalence of depression in HIV-infected Africans (ranging from 8–46%) is several magnitudes greater than prevalence estimates for the general population (Myer et al., 2008; Nakimuli-Mpungu et al., 2011; Ramirez-Avila et al., 2012). Thus, within such contexts, there is an urgent need for depression screening instruments that have maximum sensitivity and specificity to detect depression among patients living with HIV, and are brief and able to be administered by lay healthcare workers (Becker and Kleinman, 2013; Freeman et al., 2005). Such work could have a substantive impact on health and overall quality of life.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Abbreviations

CASAH	Child and Adolescent Self Awareness and Health Study
PHIV	Perinatally acquired HIV
YAPHEU	Young adults perinatally HIV exposed but not infected
YAPHIV	Young adults with perinatally acquired HIV

References

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th ed, Washington DC.

Azocar F, Arean P, Miranda J, Muñoz RF, 2001. Differential item functioning in a Spanish translation of the Beck Depression Inventory. Journal of Clinical Psychology 57, 355–365. [PubMed: 11241365]

Barroso J, Leserman J, Harmon JL, Hammill B, Pence BW, 2015. Fatigue in HIV-infected people: a three-year observational study. Journal of pain and symptom management 50, 69–79. [PubMed: 25701691]

Becker AE, Kleinman A, 2013. Mental health and the global agenda. New England Journal of Medicine 369, 66–73. [PubMed: 23822778]

Bhana A, Kreniske P, Pather A, Abas MA, Mellins CA, in press. Interventions to address the mental health of adolescents and young adults living with or affected by HIV: State of the evidence. J Int AIDS Soc.

Bhatia MS, Munjal S, 2014. Prevalence of Depression in People Living with HIV/AIDS Undergoing ART and Factors Associated with it. J Clin Diagn Res 8, Wc01–04.

Bing EG, Burnam MA, Longshore D, Fleishman JA, Sherbourne CD, London AS, Turner BJ, Eggan F, Beckman R, Vitiello B, 2001. Psychiatric disorders and drug use among human immunodeficiency virus–infected adults in the United States. Archives of general psychiatry 58, 721–728. [PubMed: 11483137]

Borowsky SJ, Rubenstein LV, Meredith LS, Camp P, Jackson-Triche M, Wells KB, 2000. Who is at risk of nondetection of mental health problems in primary care? Journal of general internal medicine 15, 381–388. [PubMed: 10886472]

Brown C, Schulberg HC, Madonia MJ, 1996. Clinical presentations of major depression by African Americans and whites in primary medical care practice. Journal of affective disorders 41, 181–191. [PubMed: 8988450]

Calonge N, Petitti DB, DeWitt TG, Dietrich AJ, Gordis L, Gregory KD, Harris R, Isham G, LeFevre ML, Leipzig RM, 2009. Screening for depression in adults: US Preventive Services Task Force recommendation statement. Annals of internal medicine 151, 784–792. [PubMed: 19949144]

Centers for Disease Control and Prevention, 2020. HIV Surveillance Report, 2018 (Updated).

Collins PY, Insel TR, Chockalingam A, Daar A, Maddox YT, 2013. Grand challenges in global mental health: integration in research, policy, and practice. PLoS Med 10, e1001434. [PubMed: 23637578]

Das AK, Olfson M, McCurtis HL, Weissman MM, 2006. Depression in African Americans: breaking barriers to detection and treatment: community-based studies tend to ignore high-risk groups of African Americans. Journal of Family Practice 55, 30–40. [PubMed: 16388764]

Do AN, Rosenberg ES, Sullivan PS, Beer L, Strine TW, Schulden JD, Fagan JL, Freedman MS, Skarbinski J, 2014. Excess burden of depression among HIV-infected persons receiving medical care in the united states: data from the medical monitoring project and the behavioral risk factor surveillance system. PLoS One 9, e92842. [PubMed: 24663122]

Freedland KE, Lustman PJ, Carney RM, Hong BA, 1992. Underdiagnosis of depression in patients with coronary artery disease: the role of nonspecific symptoms. The International Journal of Psychiatry in Medicine 22, 221–229. [PubMed: 1487385]

Freeman MC, Patel V, Collins PY, Bertolote JM, 2005. Integrating mental health in global initiatives for HIV/AIDS. The British Journal of Psychiatry 187, 1–3. [PubMed: 15994563]

Harrison DL, Miller MJ, Schmitt MR, Touchet BK, 2010. Variations in the probability of depression screening at community-based physician practice visits. Primary care companion to the Journal of clinical psychiatry 12.

Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL, 2006. Using the patient health questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. Journal of general internal medicine 21, 547–552. [PubMed: 16808734]

James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Abbastabar H, Abd-Allah F, Abdela J, Abdelalim A, 2018. Global, regional, and national incidence, prevalence, and years lived with

disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 392, 1789–1858.

Kalichman SC, Rompa D, Cage M, 2000. Distinguishing between overlapping somatic symptoms of depression and HIV disease in people living with HIV-AIDS. J Nerv Ment Dis 188, 662–670. [PubMed: 11048815]

Kang E, Mellins CA, Dolezal C, Elkington KS, Abrams EJ, 2011. DISADVANTAGED NEIGHBORHOOD INFLUENCES ON DEPRESSION AND ANXIETY IN YOUTH WITH PERINATALLY ACQUIRED HUMAN IMMUNODEFICIENCY VIRUS: HOW LIFE STRESSORS MATTER. J Community Psychol 39, 956–971. [PubMed: 23472046]

Kang E, Mellins CA, Kim W, Dolezal C, Kindler C, Leu CS, Abrams EJ, 2021. Navigating Stigma Trajectory and Mental Health Among Young Adults Living with Perinatal HIV in New York City. AIDS Behav.

Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE, 2005. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Archives of general psychiatry 62, 593–602. [PubMed: 15939837]

Kim IJ, Ge X, Brody GH, Conger RD, Gibbons FX, Simons RL, 2003. Parenting behaviors and the occurrence and co-occurrence of depressive symptoms and conduct problems among african american children. Journal of Family Psychology 17, 571. [PubMed: 14640806]

Kirmayer LJ, 2001. Cultural variations in the clinical presentation of depression and anxiety: implications for diagnosis and treatment. Journal of clinical psychiatry 62, 22–30.

Kroenke K, Spitzer RL, 2002. The PHQ-9: a new depression diagnostic and severity measure. Psychiatric annals 32, 509–515.

Kroenke K, Spitzer RL, Williams JB, 2001. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 16, 606–613. [PubMed: 11556941]

Kroenke K, Spitzer RL, Williams JB, 2003. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 41, 1284–1292. [PubMed: 14583691]

Lopes M, Olfson M, Rabkin J, Hasin DS, Alegría AA, Lin K-H, Grant BF, Blanco C, 2011. Gender, HIV status, and psychiatric disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. The Journal of clinical psychiatry 73, 384–391. [PubMed: 22053858]

Löwe B, Kroenke K, Gräfe K, 2005. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). Journal of psychosomatic research 58, 163–171. [PubMed: 15820844]

Manea L, Gilbody S, McMillan D, 2012. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. Cmaj 184, E191–196. [PubMed: 22184363]

Margaretten ME, Katz P, Schmajuk G, Yelin E, 2013. Missed opportunities for depression screening in patients with arthritis in the United States. Journal of general internal medicine 28, 1637–1642. [PubMed: 23897128]

McGoey ST, Huang KE, Palmes GK, 2013. Low depression screening rates in US ambulatory care. Psychiatric Services 64, 1068–1068.

Mellins CA, Brackis-Cott E, Leu CS, Elkington KS, Dolezal C, Wiznia A, McKay M, Bamji M, Abrams EJ, 2009. Rates and types of psychiatric disorders in perinatally human immunodeficiency virus-infected youth and seroreverters. Journal of Child Psychology and Psychiatry 50, 1131–1138. [PubMed: 19298479]

Mellins CA, Elkington KS, Leu C-S, Santamaria EK, Dolezal C, Wiznia A, Bamji M, Mckay MM, Abrams EJ, 2012. Prevalence and change in psychiatric disorders among perinatally HIV-infected and HIV-exposed youth. AIDS care 24, 953–962. [PubMed: 22519762]

Mellins CA, Malee KM, 2013. Understanding the mental health of youth living with perinatal HIV infection: lessons learned and current challenges. J Int AIDS Soc 16, 18593.

Monahan PO, Shacham E, Reece M, Kroenke K, Ong'or WO, Omollo O, Yebei VN, Ojwang C, 2009. Validity/reliability of PHQ-9 and PHQ-2 depression scales among adults living with HIV/AIDS in western Kenya. J Gen Intern Med 24, 189–197. [PubMed: 19031037]

Myer L, Smit J, Roux LL, Parker S, Stein DJ, Seedat S, 2008. Common mental disorders among HIV-infected individuals in South Africa: prevalence, predictors, and validation of brief psychiatric rating scales. AIDS patient care and STDs 22, 147–158. [PubMed: 18260806]

Nakimuli-Mpungu E, Musisi S, Katabira E, Nachega J, Bass J, 2011. Prevalence and factors associated with depressive disorders in an HIV+ rural patient population in southern Uganda. J Affect Disord 135, 160–167. [PubMed: 21851986]

New York City Department of Health and Mental Hygiene, 2007. Pediatric and adolescent HIV/AIDS surveillance update New York City: Semiannual report, December 2007. Department of Heath and Mental Hygiene, New York City

Nguyen N, Choi CJ, Robbins R, Korich R, Raymond J, Dolezal C, Leu C-S, Wiznia A, Abrams EJ, Mellins CA, 2020. Psychiatric trajectories across adolescence in perinatally HIV-exposed youth: the role of HIV infection and associations with viral load. Aids 34, 1205–1215. [PubMed: 32287067]

Owe-Larsson M, Säll L, Salamon E, Allgulander C, 2009. HIV infection and psychiatric illness. African journal of psychiatry 12.

Patterson AL, Morasco BJ, Fuller BE, Indest DW, Loftis JM, Hauser P, 2011. Screening for depression in patients with hepatitis C using the Beck Depression Inventory-II: do somatic symptoms compromise validity? General Hospital Psychiatry 33, 354–362. [PubMed: 21762832]

Posner SF, Stewart AL, Marín G, Pérez-Stable J, E, 2001. Factor variability of the center for epidemiological studies depression scale (CES-D) among urban latinos. Ethnicity & Health 6, 137–144. [PubMed: 11488294]

Ramirez-Avila L, Regan S, Giddy J, Chetty S, Ross D, Katz JN, Freedberg KA, Walensky RP, Losina E, Bassett IV, 2012. Depressive symptoms and their impact on health-seeking behaviors in newly-diagnosed HIV-infected patients in Durban, South Africa. AIDS Behav 16, 2226–2235. [PubMed: 22451351]

Remien RH, Stirratt MJ, Nguyen N, Robbins RN, Pala AN, Mellins CA, 2019. Mental health and HIV/AIDS: the need for an integrated response. AIDS (London, England) 33, 1411. [PubMed: 30950883]

Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME, 2000. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. J Am Acad Child Adolesc Psychiatry 39, 28–38. [PubMed: 10638065]

Simoni JM, Safren SA, Manhart LE, Lyda K, Grossman CI, Rao D, Mimiaga MJ, Wong FY, Catz SL, Blank MB, 2011. Challenges in addressing depression in HIV research: assessment, cultural context, and methods. AIDS Behav 15, 376–388. [PubMed: 21046221]

Skaer TL, Sclar DA, Robison LM, Galin RS, 2000. Trends in the rate of depressive illness and use of antidepressant pharmacotherapy by ethnicity/race: an assessment of office-based visits in the United States, 1992–1997. Clinical Therapeutics 22, 1575–1589. [PubMed: 11192148]

Stockdale SE, Lagomasino IT, Siddique J, McGuire T, Miranda J, 2008. Racial and ethnic disparities in detection and treatment of depression and anxiety among psychiatric and primary health care visits, 1995–2005. Medical care 46, 668. [PubMed: 18580385]

Tomlinson M, Rudan I, Saxena S, Swartz L, Tsai AC, Patel V, 2009. Setting priorities for global mental health research. Bulletin of the World Health Organization 87, 438–446. [PubMed: 19565122]

UNAIDS, 2021. AIDSinfo.

United Nations Children's Fund (UNICEF), 2017. Statistical Tables. The state of the World's children.

Uthman OA, Magidson JF, Safren SA, Nachega JB, 2014. Depression and adherence to antiretroviral therapy in low-, middle-and high-income countries: a systematic review and meta-analysis. Current Hiv/aids Reports 11, 291–307. [PubMed: 25038748]

Wagner GJ, Ghosh-Dastidar B, Mukasa B, Linnemayr S, 2020. Changes in ART adherence relate to changes in depression as Well! Evidence for the Bi-directional longitudinal relationship between depression and ART adherence from a prospective study of HIV clients in Uganda. AIDS Behav 24, 1816–1824. [PubMed: 31813077]

Yeh M, McCabe K, Hough RL, Dupuis D, Hazen A, 2003. Racial/ethnic differences in parental endorsement of barriers to mental health services for youth. Mental health services research 5, 65–77. [PubMed: 12801070]
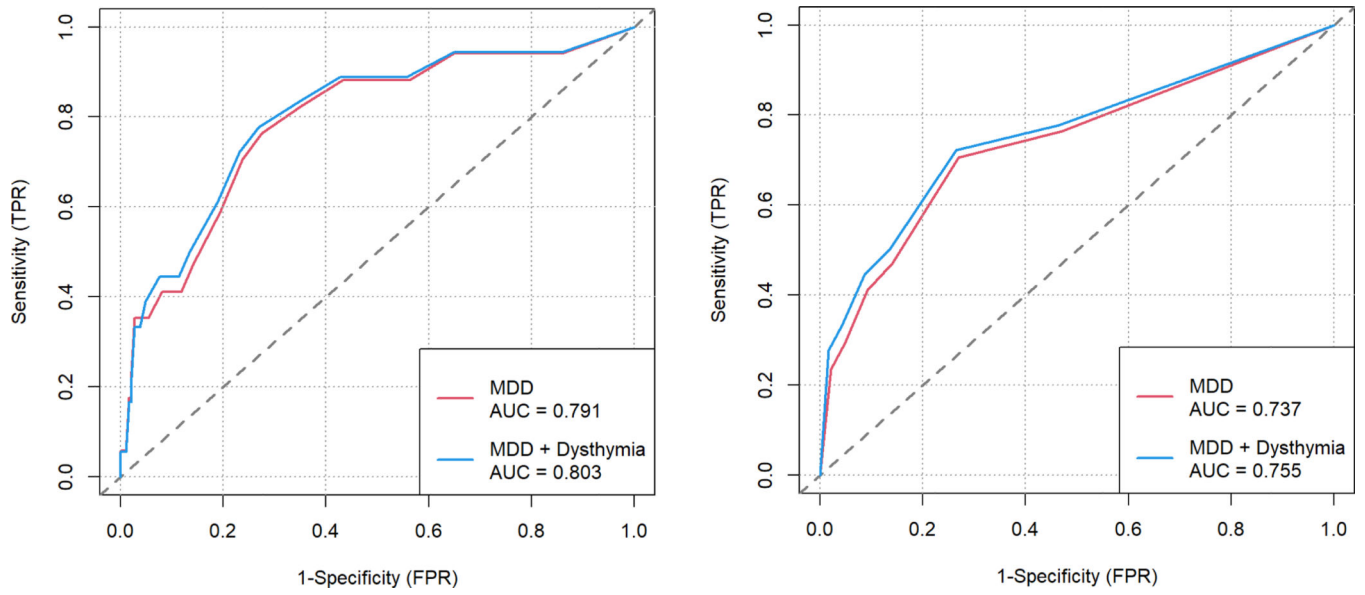
Figure 1.
ROC curves using MDD or MDD + Dysthymia as the binary classifier when considering **a)**
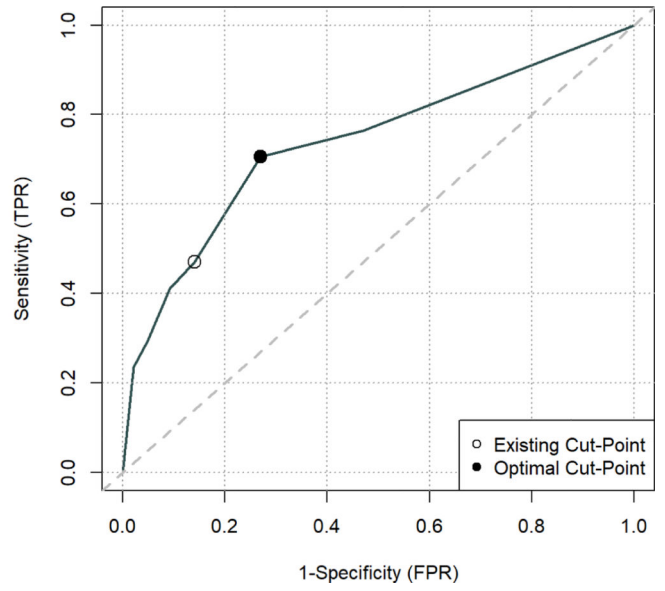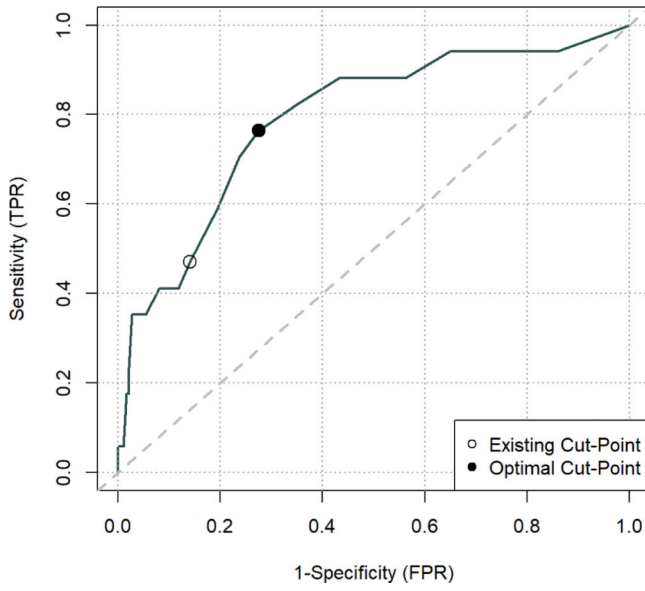PHQ-9 and **b)** PHQ-2.

Figure 2.
ROC curves with existing and optimal MDD cut-points for **a)** PHQ-9 and **b)** PHQ-2.

Table 1.

Comparison of sensitivity and specificity for depression diagnoses at the existing and optimal cut-points on the PHQ-9 and PHQ-2 in a sample of YAPHIV and YAPHEU (n=203).

| PHQ Form | DISC Diagnosis | Score | Existing Cut-Point | | Score | Optimal Cut-Point | | AUC (95% CI) |
			Sensitivity (95%CI)	Specificity (95%CI)		Sensitivity (95%CI)	Specificity (95%CI)	
PHQ-9	MDD	10	0.47 (0.23,0.72)	0.86 (0.80, 0.91)	7	0.76 (0.50,0.93)	0.72 (0.65, 0.79)	0.79 (0.66, 0.92)
	MDD + Dysthymia		0.50 (0.26,0.74)	0.86 (0.81, 0.91)		0.78 (0.52,0.94)	0.73 (0.66, 0.79)	0.80 (0.68, 0.93)
PHQ-2	MDD	3	0.47 (0.23,0.72)	0.86 (0.81,0.91)	2	0.71 (0.44,0.90)	0.73 (0.66,0.79)	0.74 (0.60, 0.88)
	MDD + Dysthymia		0.50 (0.26,0.74)	0.86 (0.81,0.91)		0.72 (0.47, 0.90)	0.74 (0.67, 0.80)	0.76 (0.62, 0.89)

Table 2.

Comparison of sensitivity and specificity for depression diagnoses at the existing and optimal cut-points on the PHQ-9 and PHQ-2 in a sample of YAPHIV (n=125).

PHQ Form	DISC Diagnosis	AUC (95%CI)	Existing Cut-Point			Optimal Cut-Point		
			Score	Sensitivity (95%CI)	Specificity (95%CI)	Score	Sensitivity (95%CI)	Specificity (95%CI)
PHQ-9		0.67 (0.58,0.75)	10	0.00 (0, 45.9)	0.85 (0.77, 0.91)	5	0.83 (0.36, 1.00)	0.55 (0.45, 0.64)
PHQ-2	MDD (n$^+$=6)*	0.57 (0.48, 0.66)	3	0.00 (0, 45.9)	0.82 (0.74, 0.89)	2	0.50 (0.12, 0.88)	0.70 (0.61, 78)

*. n$^+$ is DISC MDD=1/Yes

Table 3.

Comparison of sensitivity and specificity for depression diagnoses at the existing and optimal cut-points on the PHQ-9 and PHQ-2 in a sample of YAPHEU (n=77).

| PHQ Form | DISC Diagnosis | AUC (95%CI) | Existing Cut-Point | | | | Optimal Cut-Point | | |
			Score	Sensitivity (95%CI)	Specificity (95%CI)	Score	Sensitivity (95%CI)	Specificity (95%CI)
PHQ-9		0.86 (0.77, 0.93)	10	0.73 (0.39, 0.94)	0.88 (0.78, 0.95)	7	0.91 (58.7, 99.8)	0.74 (62.0, 84.2)
PHQ-2	MDD (n[+]=11) [*]	0.84 (0.73, 0.91)	3	0.73 (0.39, 0.94)	0.92 (0.83, 0.98)	3	0.73 (0.39, 0.94)	0.92 (0.83, 0.98)

[*]
n[+] is DISC MDD=1/Yes