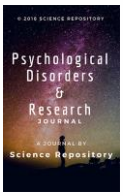

Available online at www.sciencerepository.org

Science Repository

Research Article

Behavioural Health Screening and Service Use in a Statewide Sample of Medicaid-Eligible Pediatric Outpatients

J. Michael Murphy[1,2*], Alexa Riobueno-Naylor[3], Haregnesh Haile[4], Juliana M. Holcomb[1], Anamika Dutta[1], Rao Kandru[5], Amy M. Shui[6], Hang Lee[6] and Michael S. Jellinek[1,2,7]

[1]*Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA*

[2]*Department of Psychiatry, Harvard Medical School, Boston, Massachusetts, USA*

[3]*Department of Counseling, Developmental, and Educational Psychology, Lynch School of Education and Human Development, Boston College, Boston, Massachusetts, USA*

[4]*Department of Psychology, The Catholic University of America, Washington, D.C., USA*

[5]*Department of Business Intelligence, AllWays Health Partners, Boston, Massachusetts, USA*

[6]*Biostatistics Center, Massachusetts General Hospital, Boston, Massachusetts, USA*

[7]*Department of Pediatrics, Harvard Medical School, Boston, Massachusetts, USA*

ARTICLE INFO

Article history:
Received: 7 December, 2020
Accepted: 21 December, 2020
Published: 31 December, 2020

Keywords:
Behavioural health screening
early and periodic screening
diagnosis and treatment
EPSDT
the children's behavioural health initiative
CBHI
pediatric primary care

ABSTRACT

Objective: This study explored site-level rates of behavioural health (BH) screening, positive screening, and BH service use in statewide data from the largest Medicaid Health Maintenance Organization in Massachusetts.

Methods: Screening rates at annual well-child visits (WCVs) were assessed across 908 sites (practices) for patients ages 4-17 between 2014-2018. The primary analytic sample included WCVs in 2016 (N=76,752) and was restricted to sites with at least 30 WCVs (N=304 sites; N=72,842 patients). Use of BH services was assessed six months before and after the index WCV.

Results: The mean WCV screening rate across the analytical sample was 71.2% (SD=31.3; range=0.0-100.0%) and the mean positive screening rate was 7.2% (SD=12.7, range=0-100%). Using intra-class correlations, small, but meaningful differences, were found between sites in rates of overall (r=0.38; 95% CI=0.25-0.50) and positive (r=0.10; 95% CI=0.00-0.29) screening. Although the relationship between a site's rate of screening and the rate of BH treatment failed to reach statistical significance, there was a non-significant correlation (r=0.08, p=0.17) in the predicted direction and sites that screened at or above the mean screening rate (71.9% of their WCVs) were significantly more likely to have 6.0% or more of their patients receive subsequent BH services than were sites with screening rates below 71.9% (67.2% vs. 51.2%, p<.05).

Conclusion: The current study documented a high level of continued compliance with the statewide mandate for routine psychosocial screening after more than a decade, although there were some relatively small decreases in compliance in recent years. The study also found that there were significant differences between sites in rates of BH screening and positive screening and a relationship between site-level rates of screening and service use.

Introduction

Screening for psychosocial problems has long been recommended by national standard-setting organizations as a routine part of pediatric care. The original guidelines for the U.S. Medicaid program specified that mental health screening was a required part of all well-child visits (WCVs) [1]. This guideline, which was a component of the Early and Periodic Screening, Diagnosis and Treatment (EPSDT) program, was

[*]*Correspondence to: J. Michael Murphy, Ed.D., Department of Psychiatry, Child Psychiatry Service, Massachusetts General Hospital, Yawkey 6A, Boston, Massachusetts 02114, USA; Tel: 6175103927; Fax: 6173714802; E-mail: mmurphy6@partners.org*

http://dx.doi.org/10.31487/j.PDR.2020.03.04

reaffirmed in the Medicaid Omnibus Budget Reconciliation Act of 1989 [2-4]. In 1990, the American Academy of Pediatrics (AAP) created Bright Futures as a resource for evidence-based best practices in pediatric primary care which included psychosocial screening recommendations using validated screening measures [5]. More comprehensive screening recommendations have followed over the past twenty-five years [6-8].

Early investigations demonstrated that behavioural health (BH) screening rates were low among Medicaid-insured pediatric populations [9]. In 2006, a class action suit ruling in Massachusetts (Rosie D. vs. Patrick) found that the state was out of compliance with EPSDT's BH standards which led to a court-ordered remedy - the Children's Behavioural Health Initiative (CBHI) [10-12]. CBHI provides guidelines for practices, clinician training, and an expanded array of BH services [12]. As a part of CBHI, the state required that all Medicaid-insured youth receive a brief validated psychosocial screening questionnaire at their WCV. Each screen is reimbursed, and providers are required to bill for each screen and enter a modifier code to indicate whether the screen was positive or negative, so compliance can be monitored.

Prior to CBHI, only 4.0% of Medicaid-eligible pediatric patients had evidence of a standardized BH screen at the annual WCV [13]. However, after implementation of the CBHI, rates of screening rose to 53.6% in year two and 73.9% in year five of the program [13]. Studies based on statewide billing data from the early years of the program provided evidence of its positive impact on psychosocial outcomes. One 2014 study found that screening was associated with the identification of patients at risk for psychosocial problems who were not previously in BH treatment and another found that having a positive screen increased a patient's likelihood of receiving BH services six months after screening [14, 15]. A third study which compared rates of screening and BH service use in Massachusetts and California found that rates of outpatient BH therapy increased in Massachusetts post-CBHI in comparison to California over the same period of time [16]. These findings, among others which have demonstrated the feasibility and impact of large scale psychosocial screening, were key factors in the National Quality Forum's endorsement of routine psychosocial screening in pediatrics as a national standard for quality care [17].

As of the current study, there have been no studies of statewide samples after the first few years of the program nor of statewide samples that analyse screening outcomes on a clinic- or practice-level. Since one important characteristic of adherence to quality standards is the demonstration of site-specific effects on patient outcomes, the current study explored these effects in a more recent sample, addressing the following research questions using both patient and site-level screening and BH service use data from the program's seventh to eleventh years (2014-2018):

i. Have rates of statewide BH screening and positive screening increased or decreased since the early years of CBHI?
ii. Is there a relationship between a site's screening rate and the rate of BH service utilization?
iii. What is the relationship between patient demographics, BH screening outcomes, and BH service use?

Methods

I Design

Administrative claims data were collected for pediatric patients who completed a routine WCV between 2014-2018 and were enrolled in Neighbourhood Health Plan (NHP; now AllWays Health Partners), the largest of the six Massachusetts Medicaid health maintenance organizations (HMOs). Data included billing claims for WCVs, BH screens, BH screen modifiers, BH service utilization records, and patient-level demographics (age, sex, WCV location).

II Study Sample

A total of 287,042 pediatric patients (4.00-17.99 years) presented for a WCV between 2014-2018 across 908 sites statewide. The sample was restricted to include the patients' first WCV of each year. We focused on a single year (2016) that had an index WCV screen and whose immediate prior and subsequent years had available BH service data. Therefore, the primary analytical sample included patients seen for a WCV in 2016 and used BH service data from as early as 2015 and as late as 2017 to assess BH service use six months before and after the 2016 WCV visit. Within the small number of patients with multiple WCVs in 2016, only the patient's first WCV was considered to assess BH service use six months before and after the patient's first 2016 WCV. In order to conduct meaningful site-level analyses, we further restricted the analytical sample to only sites with at least 30 WCVs to allow for appropriately-sized units of analysis, creating a sample of 304 clinics [18].

III Data

A WCV was identified for study inclusion via the following billing CPT codes: 99384, 99385, 99394, or 99395. BH screens were identified through 96110 or 96127 CPT codes which indicate that a screen had been administered and the result of the BH screen was identified through modifier codes. The U1, U3, or U5 modifier codes indicate that no BH need was identified (negative screen) and the U2, U4, and U6 codes indicate a BH need was identified (positive screen). A recent study suggests that BH claims data has a high concordance rate (94.3%) with the presence of a completed BH screen in the chart and that BH claim modifiers correspond to the screen's outcome according to scored screening data (72.9% agreement between BH risk modifier code and risk is shown on screening data) [19].

Although CPT billing codes make it possible to track the frequency and result of screens, they do not indicate which screen was completed. However, in the Massachusetts Medicaid program, only eight screens were approved for use in 2016. These included the Pediatric Symptom Checklist (PSC; parent-and youth-report measures in both the 35- and 17-item versions), the Parent's Evaluation of Developmental Status (PEDS), the Modified Checklist for Autism in Toddlers (M-CHAT), the Ages and Stages Questionnaire (ASQ), the Brief Infant Toddler Social Emotional Assessment (BITSEA), the CRAFFT Screening Test, the Patient Health Questionnaire-9 (PHQ-9), and the Strengths and Difficulties Questionnaire (SDQ) [13]. A 2016 study found that the PEDS, PSC, and MCHAT were the most frequently used pediatric screening tools in an evaluation of BH screening in Massachusetts in

2010 and 2012, with the PSC accounting for 80.0% of all screens in the 4-17-year-old age range [13].

BH service use six months before and after the patient's WCV was evidenced by 90791-90882 CPT codes (psychiatric diagnostic evaluation, individual therapy, family psychotherapy, group psychotherapy, pharmacologic management, transcranial magnetic stimulation, electroconvulsive therapy, individual psychotherapy with biofeedback, hypnotherapy, or environmental intervention for medical management purposes on a psychiatric patient's behalf) and 99201-99420 psychiatric evaluation and management codes.

IV Data Analysis

Analyses were performed using SPSS version 24 and R version 3.6.1. All hypothesis tests performed were two-sided, and significance was set at 0.05. Chi-square tests were used to compare rates of categorical BH screening, positive screening, and service use across sites. The Cochran-Armitage test was utilized to assess for directional trends in categorical BH screening and positive BH screening across five years (2014-2018). The relationship between site-level BH screening rates and rates of subsequent BH service use were assessed both continuously (Pearson bivariate correlation) and categorically (chi-square test). Rates of BH screening and subsequent BH service use were dichotomized according to the following criteria: BH screening was recoded as falling below or above the mean WCV screening rate of 71.99% in 2016 according to CBHI data reports, and BH service use was recoded as falling above or below 6.0%, which is the approximate mean percent of Medicaid-eligible pediatric patients who utilized CBHI services between 2015-2017 [20, 21].

Binary logistic regression was used to explore which variables predicted an increased likelihood of BH service utilization six months post-WCV, including prior BH service use (yes/no), completion of a BH screen (yes/no), result of the screen (positive/negative), patient sex (male/female), and patient age (child: 4.00-12.99 years/adolescent: 13.00-17.99 years). Overall and positive screening rate reliability across sites was assessed using the intra-class correlation coefficient (ICC), which assesses the proportion of total variance in screening rate that is accounted for by site variation. A higher ICC implies that the between-site variation (signal) is higher than the within-site variation (noise) [22, 23]. The R version 3.6.1 package "ICCbin" was used to compute Fleiss-Cuzick estimates of the ICC and 95% Confidence Intervals [22, 23].

Results

I BH Screening: 2014-2018

Rates of BH screening were assessed from January 1st, 2014, to August 1st, 2018. The sample of 287,042 WCVs consisted of patients with WCVs in any of the five years. For patients with multiple WCVs in a year, only the first WCV was included. The mean WCV screening rate across years and sites was 76.1%. Rates of BH screening significantly differed across the years, with a rate of 77.2% in 2015 (total N of WCV: 66,939), 75.4% in 2016 (total N of WCV: 76,022), 73.6% in 2017 (total N of WCV: 74,260), and 71.5% in January-August of 2018 (total N of WCV: 17,663; χ^2 = 1,250.16, df=4, p<.001). BH screening rates from 2014 to 2018 showed a pattern of significant decrease (Z=32.76, df=4, p<.001). Rates of positive screening, however, significantly increased from 2014 to 2018, with a mean positive screening rate of 7.3% in 2014, 7.5% in 2015, 7.6% in 2016, 11.2% in 2017, 11.9% in 2018, and a mean of 8.7% across all years. Rates differed significantly across the years (χ^2=894.79, df=4, p<.001), and positive BH screening rates by year between 2014 to 2018 followed a significant, increasing trend (Z= -25.08, df=4, p<.001).

II BH Screening: 2016

The 2016 study sample included 76,022 WCVs from 664 sites. Although only 304 sites completed at least 30 WCVs in 2016, this included most of the cases (N=72,842 WCVs). A mean of 240 WCVs (median = 125), with a range of 30-1,880 WCVs, were completed per site. Overall screening rates ranged from 0.0-100.0%, with a mean site screening rate of 75.4%. A majority of sites (n=197; 73.9%) had a screening rate above 70.0%. A majority (76.1%) of patients were below the age of 13.00 and 51.0% were male (Table 1). The ICC for screening rates or the proportion of total variance in screening rates accounted for by site variation in the 304 sites was 0.38 (95% confidence interval [CI]=0.25-0.50). Given that this proportion is ≥10.0%, the observed between-site screening rate differences can be considered meaningful [22, 23]. The percent of cases that screened positive was 7.5% overall, with a range of 0-100%. The ICC for positive screening rates in the 245 sites that also had screening results available was 0.10 (95% CI=0.00-0.29). The proportion of total variance in screening rates accounted for by site variation was 10.0%, just within the lower bound of what is considered a meaningful between-site variation.

Table 1: Patient demographics: Analytical 2016 Sample (N=72,842 patients; N=304 sites).

Patient Age Range	***n***	***%***
Child (4.0-12.9)	55,424	76.1
Adolescent (13.0-17.9)	17,418	23.9
Patient Sex		
Male	37,161	51.0
Female	35,681	49.0
Completed BH Screen		
Yes	55,759	76.5
No	17,083	23.5
Volume of WCV at Site		
30-125 WCVs	10,087	13.8
126-1880 WCVs	62,755	86.2
BH Risk (n=55,738)		
Positive Screen	4,165	7.5
Negative Screen	51,573	92.5

III Relationship Between Site-Level BH Screening and Service Use

To explore the relationship between site-level overall screening, positive screening, and BH service utilization rates in the six months following the BH screen, we first ran a correlation between site-level BH screening rates and rates of subsequent BH service use. A positive but statistically non-significant relationship between a site's BH screening rate and the rate of subsequent BH service use was found (Pearson's r=0.08, p=0.17). However, when overall site screening rates were dichotomized to contrast sites falling below (n=108 sites) with those falling above (n=196 sites) the mean screening rate of 71.9%, (as determined by CBHI) and when subsequent BH service use was dichotomized as falling below (n=45 sites) or above (n=259 sites) 6.0% (the mean rate of BH treatment in CBHI), sites that screened above 71.9% were significantly more likely to have BH treatment rates of 6.0% or above when compared to sites with less than 71.9% screening (88.3% vs. 79.4%; χ^2=4.34; df=1, p<.05), suggesting a positive relationship between BH screening rates and rates of subsequent BH service use at the site-level (Table 2).

Table 2: Categorical screening rate and subsequent BH services (N=304 sites).

	Percent Subsequent BH Service Use			
	0.0-5.99%		**6.00-30.30%**	
Percent Overall Screening Rate	***n***	***%***	***n***	***%***
0.0-71.98%	23	21.3	85	78.7
71.99-100.0%	22	11.2	174	88.8 *
Total N	45	14.8	259	85.2

BH = Behavioural Health; *χ^2= 5.60; df= 1, p<.05.

Since there was a wide range in the volume of WCVs provided across the 304 sites (N=30-1,880 WCVs), we also assessed the relationship between the number of WCVs per site and BH screening and service use. For the following analyses, the number of WCVs per site was dichotomized using a median split: sites with higher volumes of WCVs (n=153 sites with 126 to 1,880 WCVs; n=10,087 patients) and sites with smaller volumes of WCVs (n=151 sites with 30-125 WCVs; n=62,755 patients). Larger volume sites billed for screens for a significantly larger proportion of patients (78.0%) compared to smaller volume sites (67.6%; χ^2=518.50, df=1, p<.001). Prior BH service use was significantly higher (9.6%) in sites with a larger volume of WCVs compared to prior BH use in sites with a smaller volume of WCVs (8.9%; χ^2=5.20, df=1, p<.05).

IV Patient-Level Predictors of BH Service Use

A total of 9.5% of patients had evidence of previous BH service use, 10.8% had evidence of subsequent BH service use, and 7.3% had both. Of the 55,759 patients who received a BH screen, 7.5% (n=4,165) screened positively. Of these 4,165 positively screened patients, 23.9% (n=995) had prior and 28.7% (n=1,197) had subsequent BH service use. Prior BH service use significantly predicted subsequent BH service use: patients with prior BH service use (n=6,950) were significantly more likely (76.8% vs. 3.9%) to receive subsequent BH services compared to patients without evidence of prior service use (odds ratio [OR]=82.83, 95% CI=77.35-88.69, p<.001). The site-level rate of BH screening did not significantly predict site-level rates of BH service use.

For patients with a completed BH screen (n=55,759), scoring positively on the screen significantly predicted subsequent BH service use. Patients who scored positively on a BH screen were nearly four times more likely to receive BH services compared to those who did not score at-risk (28.7% vs. 9.3%; OR=3.95, 95% CI=3.67-4.25, p<.001) (Table 3). When assessing only cases with a screen and no previous BH service use (n=50,447), scoring positively on a BH also significantly predicted subsequent BH service use, with 12.0% of those who scored positively going on to receive subsequent services versus 3.2% of those who screened negatively (OR= 4.10, 95% CI=3.64-4.62, p<.001). In the full sample, age and sex were also significant predictors of BH service use post-screen. Male patients were significantly more likely than females (12.0% vs. 9.5%) to receive subsequent BH services (OR=1.30, 95% CI=1.13-1.06, p<.001). Adolescents were also significantly more likely (14.2% vs. 9.7%) to receive subsequent treatment post-WCV compared to children (OR=1.57, 95% CI=1.48-1.67, p<.001).

Table 3: Screening outcome and subsequent BH services (N=55,738 patients; N=291 sites).

	Subsequent BH Service Use			
	Yes		**No**	
Screen	***n***	***%***	***n***	***%***
Negative	4,775	9.3	46,798	90.7
Positive	2,968	71.3	1,197	28.7 ***
Total N	5,972	10.7	49,766	89.3

BH = Behavioural Health; ***χ^2= 1,528.79; df= 1, p<.001.

Discussion

The current study assessed compliance six to eleven years after the start of the CBHI screening mandate in the largest Medicaid HMO in Massachusetts by evaluating screening rates and BH service billing using site- and patient-level data. The study found a wide range of overall and positive screening rates across sites, and an ICC analysis found meaningful site by site differences in both rates. Categorical assessments of site-level screening rates and BH service use demonstrated that practices with WCV screening rates higher than the mean state-level screening rate had a higher proportion of subsequent BH service use. Additionally, in the sample of patients with no prior BH service use, evidence of a BH screen significantly predicted BH service use following the screen. Taken together, these findings suggest that site-level differences in adherence to BH screening mandates may be related to a small but significantly greater likelihood of BH service use for pediatric patients, specifically those who screen positively and who are, therefore, presumably in greater need of services. However, results also suggest a disparity in access to BH services since the rates of BH screening and of prior and subsequent BH services varied widely across sites.

The shared goal of the court and state Medicaid authority in establishing CBHI was to increase the identification and treatment of children at an early point in the development of BH problems. The current study

provides evidence showing that even after more than a decade, in Massachusetts, that goal continues to be met, with hundreds of thousands of children screened each year and thousands receiving BH services. However, despite this decade-long success story, the current study also provided some cause for concern, with the screening rates reported in this study and in statewide data decreasing in the last few years while positive screening rates were increasing [21]. Therefore, there is a need for more research related to the following questions: Why are screening rates decreasing and why are rates of positive screens increasing? What should the clinical or administrative response be if no more than one-quarter of positively screened children receive treatment? Should BH screening scores in individual patients be tracked to assess their progress (or lack thereof) over time? Are there specific barriers to receiving treatment in this population? What will be the impact of COVID-19 on screening and service use? Much of the information needed to answer these questions is available now and analysing it could serve to increase the efforts to continuously improve the quality of the CBHI screening program.

Highlights

i. Although on a statewide basis more than 70.0% of Medicaid-insured pediatric patients had evidence of a completed BH screen more than a decade after the start of the Children's Behavioural Health Initiative (CBHI), rates of behavioural health (BH) screening and positive screening were found to significantly differ across sites.

ii. Patients with a positive BH screen were three times more likely to receive BH services in the six months following their screen compared to those who did not score at-risk, further supporting previous evidence suggesting that BH screening informs BH service referrals for children in need.

iii. Since small decreases in screening rates were noted across the study period, a quality assurance program might shed light on new challenges facing CBHI compliance.

Acknowledgements

The authors wish to thank the Fuss Family Fund for their generous support which paid for a portion of the salaries of Dr. Murphy, Ms. Riobueno-Naylor, Ms. Haile, Ms. Holcomb, and Ms. Dutta.

Conflicts of Interest

None.

Ethical Approval

The study was approved by the Partners Institutional Review Board as non-human subjects research and complied with HIPAA Safe Harbor requirements. A data use agreement was fully executed between the Massachusetts General Hospital research team and Neighbourhood Health Plan/AllWays Health Partners.

Disclosure

The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

1. EPSDT: An Overview (2005) New York, The Commonwealth Fund.
2. Early Periodic Screening, Diagnosis, and Treatment (2018) Rockville, Health Resources & Services Administration.
3. Rosie D. and Mental Health Screening: A Case Study in Providing Mental Health Screening at the Medicaid EPSDT Visit (2010) New York: Teen Screen National Center for Mental Health Checkups at Columbia University,.
4. Federal Legislative Milestones in Medicaid and CHIP. Washington DC, MACPAC.
5. Hagan JF, Shaw JS, Duncan PM (2017) Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents [pocket guide]. 4th ed. Elk Grove Village, IL, *Am Acad Pediatr*.
6. Links to Commonly Used Screening Instruments and Tools. Washington DC. *Am Acad Pediatr*.
7. Committee on Psychosocial Aspects of Child and Family Health and Task Force on Mental Health (2009) Policy statement--The future of pediatrics: mental health competencies for pediatric primary care. *Pediatrics* 124: 410-421. [Crossref]
8. Foy JM, Kelleher KJ, Laraque D, American Academy of Pediatrics Task Force on Mental Health (2010) Enhancing pediatric mental health care: strategies for preparing a primary care practice. *Pediatrics* 125: S87-S108. [Crossref]
9. Semansky RM, Koyanagi C, Warren RV (2003) Behavioral health screening policies in Medicaid programs nationwide. *Psychiatr Serv* 54: 736-739. [Crossref]
10. Overview of the Case and Summary of the Trial in Rosie D. v. Romney. Boston, Federation for Children with Special Needs 2006.
11. Early and Periodic Screening, Diagnostic, and Treatment. Baltimore, Medicaid.gov.
12. Learn about CBHI: MassHealth Children's Behavioral Health Initiative (CBHI). Mass.gov.
13. Savageau JA, Keller D, Willis G, Kathleen Muhr, Gideon Aweh et al. (2016) Behavioral Health Screening among Massachusetts Children Receiving Medicaid. *J Pediatr* 178: 261-267. [Crossref]
14. Hacker KA, Penfold R, Arsenault L, Zhang F, Murphy M et al. (2014) Screening for behavioral health issues in children enrolled in Massachusetts Medicaid. *Pediatrics* 133: 46-54. [Crossref]
15. Hacker KA, Penfold RB, Arsenault LN, Fang Zhang, Michael Murphy et al. (2014) Behavioral health services following implementation of screening in Massachusetts Medicaid children. *Pediatrics* 134: 737-746. [Crossref]
16. Hacker K, Penfold R, Arsenault LN, Zhang F, Soumerai SB et al. The impact of the Massachusetts behavioral health child screening policy on service utilization. *Psychiatr Serv* 68: 25-32. [Crossref]
17. National Quality Forum (2018) Behavioral Health and Substance Use Fall 2017 Cycle: CDP Report. National Quality Forum.

18. Hogg RV, Tanis EA (1988) Probability and statistical inference. Ann Arbor MI, University of Michigan Publishing.
19. Boudreau AA, Naylor AR, Haile H, Holcomb JM, Lucke CM et al. (2020) How an Electronic Medical Record System Facilitates and Demonstrates Effective Psychosocial Screening in Pediatric Primary Care. *Clin Pediatr* 59: 154-162. [Crossref]
20. CBHI Data Reports. Service Utilization quarterly reports. Mass.gov.
21. CBHI Data Reports. BH Screening Cumulative Quarterly Report 2012-March 2019. Mass.gov.
22. Fleiss JL, Jack C (1979) The Reliability of Dichotomous Judgments: Unequal Numbers of Judges per Subject. *Appl Psych Meas* 4: 537-542.
23. Zou G, Donner A (2004) Confidence Interval Estimation of the Intraclass Correlation Coefficient for Binary Outcome Data. *Biometrics* 60: 807-811. [Crossref]