



MeHAF Integration Initiative: Outcomes Analysis for Clinical Implementation Grantees

Prepared for:

Maine Health Access Foundation
150 Capitol Street, Suite 4
Augusta, ME 04330

Prepared by:

John Snow, Inc.
44 Farnsworth Street
Boston, MA 02210



Table of Contents

Introduction	1
Methods.....	1
2007 Clinical Implementation Grantees	3
Amistad	3
Aroostook Mental Health Center (AMHC)	3
Community Counseling Center	4
Community Dental (CD)	5
DFD Russell Medical Centers	6
Maine-Dartmouth Family Medicine Residency (MDFMR).....	7
Pen Bay Health Care.....	8
Sacopee Valley Health Center (SVHC).....	9
Spring Harbor 2007	10
York County Community Health Care (York).....	12
2008 Clinical Implementation Grantees	12
Acadia Hospital	12
Downeast Health Services	13
Franklin Health Pediatrics (FHP)	14
Northeast Integrated Geriatrics Care	15
Spring Harbor 2008	17
Tri-County Mental Health Services (TCMHS)	18
2009 Clinical Implementation Grantees	20
Eastern Maine Medical Center (Center for Family Medicine)	20
Mercy (Fore River Clinic).....	22
Conclusion.....	24

Introduction

The Maine Health Access Foundation (MeHAF) launched its Integration Initiative in 2006 to promote patient and family-centered care by improving the integration of behavioral and physical health care services. Between 2007 and 2009, MeHAF funded three rounds of grants, including one- and two-year Planning and three-year Clinical Implementation, Data, and Systems Transformation grants. Forty-two grants were awarded and involved over 100 practices and over 150 partnering organizations.

In January 2009, John Snow, Inc. (JSI) was contracted by MeHAF to conduct a cross-site evaluation of the Clinical Implementation grants. The RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance) framework was selected to guide the evaluation (see www.re-aim.org). This report is a companion piece to the major evaluation reports submitted to MeHAF in January 2010 (Year One) and in January 2011 (Year Two). This report focuses on the “Effectiveness” domain of the evaluation framework, which concerns the outcomes related to the Integration Initiative. Very specifically, this report examines patient-level, clinical outcomes, as one way of addressing the question:

Among people who had access to MeHAF initiative-funded integrated services, what was the impact on their health?

The clinical outcomes evaluation process was developed in 2009 after the other cross-site evaluation components were in place. Outcomes data collection began in 2010. Primarily intended for grantee’s program planning purposes, the data were also incorporated into the cross-site evaluation as a descriptive snapshot of clinical outcomes at participating practices. Grantees submitted data as they were best able without specific grant resources dedicated to technical assistance to support infrastructure (e.g., registries, software) improvements for outcomes reporting. Grantees that were more successful in providing data were working with practices within their own organizations (rather than partners) because they had better access to the data. Grantees that had some prior experience with quality improvement initiatives also had better success reporting clinical outcomes data. This report examines the outcomes information collected from the 2007, 2008, and 2009 Clinical Implementation Grantees.

Methods

From November 2009 through March 2010, each implementation grantee worked with its MeHAF project officer to identify patient-level clinical indicators (outcomes measures). The choice of indicator(s) was based on finding measure(s) that were relevant to the grant project’s approach and patient population, feasible to collect, and could be measured reliably and validly. Because of the wide variance of approaches to integrated care among grantees, there was no intention that a common outcome would be selected for the cross-site evaluation or that outcomes would be aggregated across all grantees. During the spring of 2010, JSI had conference calls with each grantee to work out analytic (e.g., timing of follow-up measurement, identifying which patients to measure) and operational issues (e.g., identifying data sources, creating data formats). Data draws were completed semi-annually and sent to JSI using an Excel template provided by JSI. Data were collected at the practice site level, with the grantees responsible for gathering spreadsheets from all of their practice sites in the case of multi-

site projects. Patients who had been assessed and then referred for MeHAF integrated services were the target population for the outcomes analysis.

The same process was used in 2010 to plan for outcomes measurement for the 2008 and 2009 implementation grantees. The 2007 and 2008 grantees began outcomes data collection in 2010, whereas the 2009 grantees began in 2011. Among the 20 grantees, 15 different patient-level indicators were chosen. Seven grantees chose the PHQ-9 for depression (often pairing it with the GAD-7 for anxiety). Otherwise, measures were unique to each grantee.

Three 2007 and one 2009 clinical implementation grantees did not provide any outcomes data for their sites. The following table lists grantees and their selected outcomes measures.

Grantee Name	Outcome Indicator
<i>Grantees that submitted outcomes data - included in report</i>	
Acadia Hospital-2008	SF12v2
Aroostook Mental Health Services	OQ-45
Amistad	Body weight
*Community Counseling Center	GAF
Community Dental	Referrals
DFD Russell Medical Centers	PHQ-9
*Downeast Health Services	Treatment Plan Goals
Eastern Maine Medical Center (Center for Family Medicine)	Substance Use Screening
*Franklin Health Pediatrics	Autism Screening
*Maine-Dartmouth Family Medicine Residency	Resilience
Mercy (Fore River Family Practice)	PHQ-9, Beck/GAD-7 for DM, HBP, LBP patients; also: HbA1c (DM), blood pressure HBP), pain measure (LBP)
Pen Bay Healthcare (Mid Coast Mental Health Center)	PHQ-9, GAD-7
Northeast Integrated Geriatric Care	Sadness
Sacopee Valley Health Center	HbA1c
Spring Harber-2008	Survey
Spring Harbor-2007	PHQ-9
Tri-County 2008 and 2009	PHQ-9, GAD-7
York County Community Health Care	PHQ-9
<i>Grantees that did not submit outcomes data (excluded from report)</i>	
Acadia Hospital-2007	
Penobscot Community Health Center (Husson Capehart)	Survey
HealthReach Community Health Centers	
St. Mary's Health System	PHQ-9

* Program or outcomes data focused on the pediatric population.

Reporting practices faced two key challenges regarding outcomes data:

- Measuring patients over time after the initial assessment (aka “follow-up assessment”); and

- Extracting data from medical record or project-specific data systems that correctly identified patients who had been assessed and referred for MeHAF-funded integrated services.

Given the variety of outcomes and the variety of approaches to address the data challenges, it was not feasible to compute aggregated, project-wide outcomes results. Instead, this report contains summaries of the outcomes data for each implementation grantee. Key thematic findings are presented in the conclusion section.

Each grantee summary provides a very brief program overview and describes the number of clients/patients served by the program, cross-referencing the outcomes data with the Client Data Elements (CDE) database. The CDE database is the process indicator database used for the cross-site evaluation, containing counts of patients screened, assessed, referred and seen for mental/behavioral health needs. To keep summaries succinct and clear, detailed results are presented only when there are at least 10-20 patients to report. For this reason, data on pediatric outcomes are not reported unless the focus of the program or outcomes data was on children and young adults.

2007 Clinical Implementation Grantees

Amistad

This project provided Amistad members (those with severe and persistent mental illness) with peer patient navigator support to access health services. It also encouraged healthy eating through onsite education and provision of healthier lunch options. The outcome measure for Amistad was weight change. Weight data came from 30 Amistad members who volunteered to have their weight taken at the beginning of the project and every month thereafter for six months. Nearly one third (9/30=30%) of the members were not re-weighed. For the 22 members who were re-weighed, the average number of months between the initial and last weighing was 4.6 months (range: 2 to 6 months).

Four members whose goal was to gain weight had available initial and follow-up data. Their initial average weight was 144.5 pounds, and they gained an average of 8.25 pounds over an average follow-up time of 3.25 months. Two of the 4 gained over 5 lbs. and did so over the course of 4-6 months. The other 2 members who had less gain were followed for only 1-2 months.

Eighteen members whose goal was to lose weight had available initial and follow-up data. Their initial average weight was 266.1 pounds, and they lost an average of 9.11 pounds over an average follow-up time of 4.6 months. Nine of the 18 (50%) lost more than 5 pounds, 8 of the 18 (44%) had weights within +/- 5 lbs. of their original weight, and only 1 gained weight (7 pounds). The four members with the greatest losses (10 to 65 pounds) did so over the course of 4-6 months. One member weighed over 600 pounds and achieved a weight loss of over 10% - an excellent achievement.

Aroostook Mental Health Center (AMHC)

AMHC contracted with the Pines Family Health Center and Fish River Rural Health, both federally-qualified health centers (FQHCs), to provide on-site behavioral health/mental health (BH/MH) services. According to the CDE, during the first half of 2010, there were 96 referrals at Pines and 22 at Fish River

for BH/MH services. Unfortunately, outcomes data were only available for 15 Fish River patients and no Pines patients. No outcomes data were available for the second half of 2010.

The outcome measure was the OQ-45, which according to the scoring guide¹, is a 45-item assessment of symptom distress, interpersonal relations, and social role. The scale ranges in values from 0 to 180 and a score of ≥ 63 represents clinically-significant stress. Over time, a change of 14 points or more is considered a reliable difference. Among the Fish River patients, 73% (11 of 15) had an initial OQ-45 above the cut point, with a mean score of 89.7 (standard deviation = 39.1). Only 6 patients had follow-up scores, measured about two months after the initial score. Four of these patients experienced symptom reduction, although only one of them was a reduction of 14 points. Those with the highest scores had the greatest reductions, but two months may not be long enough time for clinically-significant change.

Fish River: Patient Change Scores on the OQ-45

Patient	Initial OQ-45 Score	Follow-up OQ-45 Score	Change
1	104	91	-13
2	55	42	-13
3	100	86	-14
4	19	25	6
5	69	73	4
6	52	45	-7

Although these results are interesting and promising, they represent only 5% of all clients (6 of 118), and none from Pines, so it is not possible to draw conclusions about the outcomes for this grantee as a whole.

Community Counseling Center

This grantee integrated BH/MH services at school-based health centers in Portland High School (HS) and Deering High School. During 2010, outcomes data were captured for all 21 Portland HS students who were referred for BH/MH services and for 19 Deering HS students.

The Global Assessment of Functioning (GAF) was the outcome measure selected by Community Counseling. The GAF represents the clinician’s subjective judgment of a patient’s overall level of psychological, social, and occupational functioning. The clinician bases this judgment on a DSM-IV assessment (Axis V). The 100-point scale is broken into 10-point increments, whereby patients who are deemed to have superior functioning are scored at the top of the range (91-100); patients with moderate symptoms are scored in the middle of the range; and patients with serious symptoms are scored at the bottom of the range (e.g., 21-30, hallucinations). A 10-point change in assessment represents a clinically-significant change in function.

¹ Lambert, M. J., Lunnen, K., Umphress, V., Hansen, N. & Burlingame, G. M. (1994). *Administration and scoring manual for the Outcome Questionnaire (OQ-45.1)*. Salt Lake City: IHC Center for Behavioral Healthcare Efficacy.

Among the 40 high school students, the average initial GAF score was 59.9 (standard deviation = 7.53, range = 46 to 70), indicating moderately serious impairment. After an average follow-up of 120 days, the average GAF score was essentially unchanged at 61.3 (standard deviation = 9.31, range = 40 to 87). Two students had clinically-significant gains in function, one improving 17 points from a baseline of 70 (mild symptoms) to 87 (absent symptoms) and the other improving 18 points (from moderate to transient symptoms). Both students were followed for about 140 days. One symptomatic student was re-evaluated only two weeks later, with a significant decrement in GAF score (mild to serious symptoms). However, the time between measures was likely not long enough for a therapeutic effect. The seeming lack of improvement for the other youth could be due in part to the complexity of the issues these youth were facing, lack of engagement in care or services, the subjective nature of the GAF, and/or the inter- or intra-rater reliability of the clinicians using it. It is not possible to determine from the data which, if any, of these factors played a role in the results.

The following table is taken from the GAF scoring guide and describes the functioning in the middle of the range where most students' scores fell. The two youth who had substantial improvements are noted in the table to portray their likely symptomatic profile.

Description of the Range of Symptoms on the GAF Scale

90 - 81	Absent or minimal symptoms (e.g. mild anxiety before an exam), good functioning in all areas, interested and involved in a wide range of activities, socially effective, generally satisfied with life, no more than everyday problems or concerns (e.g. an occasional argument with family members).	Youth #1: Follow-up score=87
80 - 71	If symptoms are present, they are transient and expectable reactions to psychosocial stressors (e.g. difficulty concentrating after family argument); no more than slight impairment in social, occupational, or school functioning (e.g. temporarily falling behind in schoolwork).	Youth #2: Follow-up score=73
71 - 61	Some mild symptoms (e.g. depressed mood and mild insomnia) OR some difficulty in social, occupational, or school functioning (e.g. occasional truancy, or theft within the household), functioning pretty well. Has some meaningful interpersonal relationships.	Youth #1: Initial score=70
60 - 51	Moderate symptoms (e.g. flat affect and circumstantial speech, occasional panic attacks) OR moderate difficulty in social, occupational, or school functioning (e.g. few friends, conflicts with peers or co-workers).	Youth #2: Initial score=55

Community Dental (CD)

The primary purpose of this project was to establish a dental health center in Rumford. Community partners Rumford Hospital Emergency Department, Tri-county Mental Health, and three primary care practices (Elsmore-Dixfield, River Valley Internal Medicine, and Swift River Family Medicine) were referral sources for the dental practice. During 2010, Community Dental tracked the number of referrals from each practice in the CDE and attempted to abstract the exam code from patient charts to elucidate the reason for referral. There were two exam codes used: *evaluation*, for a comprehensive check-up, and *limited*, for an acute need. The grantee hypothesized that people whose initial visits were for evaluation would be more likely to come back to CD for further care than people whose initial visits were for acute needs who might not be seen again.

The Rumford Hospital Emergency Department was the primary referral source among the partners, generating 63 patients during 2010 (out of 112 total referred patients, or 56%). The River Valley practice was the second greatest source, generating 30 patients (of 112, or 27%). Exam code was successfully abstracted from charts during the first half of 2010, but the data were not complete for the second half of 2010. During the first half of 2010, 85% (45 of 53) of referrals were for evaluation visits; the remaining 8 referrals were of the limited visit type.

DFD Russell Medical Centers

DFD Russell Medical Centers is an FQHC with three practice locations in central Maine. This grantee had prior experience with data collection and analysis for quality improvement, which proved very valuable for this project. This was evidenced by the facts that it was one of only two 2007 grantees able to consistently identify MeHAF-project related patients in terms of process data and outcomes data and to consistently measure outcomes over time. The outcomes measurement focused exclusively on depressed patients who were referred in either the first or third quarter of 2010. DFD Russell’s well-specified methodological approach facilitated the outcomes data analysis and provided substantive findings for patients with a diagnosis of depression. One limitation however, in the ability to describe the total impact of integrated services was that outcomes data were not collected for patients receiving BH/MH services for conditions other than depression (44%).

2010 Number of Referrals for DFD Russell Medical Centers Sites

Site	Q1 Total Patients	Q1 Depressed Patients	Q3 Total Patients	Q3 Depressed Patients
Leeds	53	28	33	14
Monmouth	40	22	37	20
Turner	87	51	38	24
Total	180	101 (56%)	108	58 (56%)

The PHQ-9 is part of routine care at each of the three medical center sites. It is completed at the start of every primary care visit and prior to behavioral health visits for depressed patients. For this analysis, the follow-up score that was dated closest to 90 days after the initial score was identified, resulting in an average number of days between the initial and follow-up score of between 50 to 62 days, depending on the site.

For all three sites combined, 132 of 159 (83%) depressed patients had both initial and follow-up scores. The average time between scores was 54.5 days. The average initial PHQ-9 score was 13.7, and it dropped an average of 5.6 points, or 41% to 8.1, a statistically-significant drop in mean PHQ-9 scores.

All three DFD Russell sites – Leeds, Monmouth, and Turner – experienced a significant drop (improvement) in mean PHQ-9 scores among their patients with major depression. At Leeds, 33 of 42 (78%) depressed patients had both initial and follow-up scores. The average initial PHQ-9 score was 14.6 and it dropped an average of 7.5 points, or 51%, to 7.1. Not only did the mean score drop, but so did the variability (the standard deviation went from 7.8 to 5.6), indicating there were fewer very high scores at follow-up. At Monmouth, 35 of 42 (83%) depressed patients had both initial and follow-up

scores. The average initial PHQ-9 score was 15.4 and dropped an average of 6.1 points, or 40%, to 9.3 with a slight increase in variability (6.4 to 7.3). At Turner, 64 of 75 (85%) depressed patients had both initial and follow-up scores. The average initial PHQ-9 score was 12.3, and it dropped an average of 4.3 points, or 35%, to 8.0, with no change in variability (standard deviation of 5.9 at both points in time). For all three sites, these were statistically significant reductions. This was determined by computing 95% confidence intervals (t-distribution) around the mean baseline and follow-up scores and finding there was no overlap between the intervals.

Another way to interpret the PHQ-9 outcomes is to consider how many patients had a 50% or more reduction in symptoms or a follow-up score less than 5 (remission). Nearly half (64/132, or 48%) achieved this level of improvement. Two thirds (67%) of Leeds, 34% of Monmouth, and 47% of Turner patients achieved either or both results.

DFD Russell Medical Centers: PHQ-9 Outcomes for Depressed Patients

Site	Statistic	Initial PHQ-9	Follow-up PHQ-9	Average # Days of Follow-up	Number (%) of Patients w/ Complete Data
Leeds	Mean	14.6	7.1	62.2 days	33 of 42 (77%)
	Standard Deviation	7.8	5.6		
	95% CI of mean	(11.3, 16.9)	(5.1, 9.1)		
Monmouth	Mean	15.4	9.3	54.8 days	35 of 42 (83%)
	Standard Deviation	6.4	7.3		
	95% CI of mean	(13.0, 17.4)	(6.8, 11.8)		
Turner	Mean	12.3	8.0	50.5 days	64 of 75 (85%)
	Standard Deviation	5.9	5.9		
	95% CI of mean	(10.8, 13.8)	(6.5, 9.5)		
Overall	Mean	13.7	8.1	54.5 days	132 of 159 (83%)
	Standard Deviation	6.6	6.2		
	95% CI of mean	(12.53, 14.81)	(7.02, 9.16)		

Maine-Dartmouth Family Medicine Residency (MDFMR)

A collaboration between Maine Dartmouth Family Medicine Residency and Bread of Life Ministries (BOLM – a homeless shelter), this project provided coordinated health and social services to families in BOLM transitional housing. Another partner, Crisis and Counseling Centers, provided a licensed clinical social worker (LCSW) on site at MDFMR for adult referrals, and the project hired a psychologist on site for pediatric/family referrals. Seventy adults and children were referred for integrated services during the first half of 2010 (85 for the entire calendar year). However, the project lacked support for data collection and analysis; thus, no outcome information was available for adult patients. There was some limited information on children and families because the psychologist had the interest and expertise to develop an assessment form for families and children seen through the grant, the Child/Family Protective Factors form, the content of which was based on his research and experience on factors related to family stress. As described in an April 2010 MDFMR report, the form was useful for informing the providers (the psychologist and the pediatrician) about family issues and strengths, creating a

mutually-developed intervention plan, tracking progress, and providing information for program evaluation.

The 15-question Child/Family Protective Factors assessment had four components, with each component having 3-4 questions: 1) services (“my family has adequate housing, transportation, health services, parent services”); 2) parental support (“I have someone understanding to talk to in general and about children, I have support in times of crisis, I have people I enjoy doing things with”); 3) parenting skills (“I understand how my child is developing, I am able to manage my child’s development, my child and I enjoy being together”); and 4) parent resilience (“understanding whether parent’s past positive and negative experiences impact on parenting, ability to manage past stresses impact on parenting, parenting is meaningful to me”). Each question was ranked by the parent on a 5-point scale (0=not true to 4 = very true; so a high score is better).

Child/Family Protective Factors assessment data were provided for 6 of the 9 families under care during the first half of 2010. Three families also had follow-up data.

The average baseline scores for the 4 components for the 6 families were higher (better) for access to services (3.2) and support (3.5) and lower for parenting skills (2.3) and resilience (2.9). For the 3 families with follow-up data, 1 family scored a perfect 4.0 at baseline on all components, so could not show improvement over time using the numerical protective factors score. The other 2 families, over the course of 3 or 8 months, felt they improved their skills and resilience (improvement in at least one question in each component) but did not improve on support or services. There is face validity to these results in that the psychologist likely could have a more direct influence on improving parenting skills and parental resilience than on access to services or quality of social support.

This assessment tool was useful to the psychologist-pediatrician collaboration. However, there was not enough data to assess the project’s clinical effectiveness more generally.

Pen Bay Health Care

There were nine integration sites (primary care with co-located behavioral health) involved in this project in 2010. Five of these sites provided integrated MH/primary care services to youth through three school based health centers (15 referrals) and two pediatric practices (62 referrals). Because of the small number of pediatric outcome records (9), most with incomplete data, they are not included in this analysis. Four of Pen Bay’s 9 integration sites provided integrated services to adults: Pen Bay Family Medicine (101 referrals), Pen Bay Internal Medicine (87 referrals), Waldoboro Family Medicine (31 referrals) and a family medicine physician in private practice (18 referrals). These sites are the focus of this analysis.

This grantee had difficulty extracting outcomes data from its computer systems. Specifically, project staff had difficulty identifying MeHAF-project related patients, as evidenced by the records in the data set for non-participating sites and missing dates of service. After data cleaning, there were 313 adult patients with baseline PHQ-9 and GAD-7 scores from the participating sites. This number is greater than the number of referrals captured in the CDE. Only 27 (9%) had at least one follow-up set of assessments

(6 patients had 2 or more follow-up assessments). It is not possible to know if the large number of patients and the low follow-up rate was due to difficulty extracting data, low retention in care, or a combination of both.

The initial adult PHQ-9 depression and GAD-7 anxiety scores indicated substantial levels of comorbidity. The average PHQ-9 score was 13.1 (standard deviation 6.7; inter-quartile range (IQR): 8 to 19) and the average GAD-7 score was 12.7 (standard deviation = 5.9; IQR: 8 to 18). Pen Bay Family Medicine had significantly higher average scores than Pen Bay Internal Medicine (average PHQ-9: 14.3 vs. 11.9; average GAD-7: 13.6 vs. 11.6). Over half (182 patients or 53%) had scores above 9 on both scales, indicating both moderate-to-severe anxiety and depression.

Initial Depression and Anxiety Levels of Adult Pen Bay Patients (n=313)

	No/Mild Anxiety (GAD-7 0-9)	Moderate/Severe Anxiety (GAD-7 >9)
No/Mild Depression (PHQ-9 0-9)	63 (20%)	35 (11%)
Moderate/Severe Depression (PHQ-9 > 9)	33 (11%)	182 (53%)

The 27 patients with follow-up data scores had, on average, higher baseline scores (PHQ-9 = 14.6, GAD-7 = 13.6). On average, their PHQ-9 score dropped 3.1 points and their GAD-7 scores dropped 2.5 points. Only 3 (11%) had a PHQ-9 score that was less than 5 at follow-up (indicating remission) and 5 (19%) had a 50% or greater reduction in either depression or anxiety symptoms. With such a small sample (and unknown reasons why these patients had follow-up data), it is not clear whether their results were representative of all patients’ experience.

Sacopee Valley Health Center (SVHC)

Sacopee Valley Health Center is a single site FQHC in western Maine. Over 400 SVHC patients were referred to the BH/MH integrated services in 2009 and over 300 in 2010. SVHC chose to focus on a relatively small, targeted segment of this group for outcomes analysis - 47 patients who were also diabetic and had high HbA1c levels. BH/MH services were offered so patients could learn to better manage their diabetes. SVHC was one of the two 2007 grantees that was able to successfully identify MeHAF project patients and track them over time, providing consistent process and outcome data. Patients who were referred to the BH /MH integrated services during the last quarter of 2009 and the first quarter of 2010 were selected. The 47 patients represent about 18% of the 258 diabetic patients who get their usual care at SVHC (meaning that they had >= 2 visits in the past year), and 13% of the 366 diabetic patients who visited SVHC at least once in 2010.

SVHC tracked the degree with which these patients engaged in behavioral health services. Among 47 diabetics, 16 (34%) fully participated in the program, whereas 10 (21%) started in the program but did not complete the program. There were 17 (36%) patients who did not engage after expressing initial interest in the program and 4 (9%) transferred out of SVHC. Care management notes indicated there

was an average of 9 telephone contacts per fully-participating patient. Among non-completers and patients who changed their minds and chose not to participate, 56% (15/27) seemed to do so implicitly by not returning messages (2-4 messages left per patient).

The outcome measure chosen was the blood test HbA1c, which is a measure of the average amount of sugar in the blood (glycated hemoglobin) over three months. The normal range is 4% to 6%, and for people with diabetes, the goal is to keep the value at less than 7%. Initial HbA1c values were equivalent across the four groups of patients, ranging from 10.3 to 10.5%. Most patients (except transfers) had two follow-up measures, taken about 3-5 months later, and then 8-10 months later. Only one patient (in the non-response group) had an HbA1c reading less than 7% during follow-up and 4 patients (2 in the participating group) had an upward trend in readings. Overall, there was a non-statistically significant trend toward lower HbA1c in all three groups that continued to receive their care at SVHC.

HbA1c Results over time among SVHC Diabetic Patients

	Group	Initial Value	Follow-up (3-5 months)	Follow-up (8-12 months)
Full Participants	Number of patients	16	16	14
	Mean (st. dev.) HbA1c	10.4 (1.49)	9.0 (1.42)	9.4 (1.86)
	95 % CI	(9.6, 11.2)	(8.2, 9.8)	(8.3, 10.5)
Noncompleters	Number of patients	10	6	7
	Mean (st. dev.) HbA1c	10.5 (1.04)	10.2 (1.30)	8.9 (1.47)
	95 % CI	(9.8, 11.2)	(8.8, 11.5)	(7.5, 10.3)
Not Engaged	Number of patients	17	10	12
	Mean (st. dev.) HbA1c	10.3 (0.84)	10.0 (1.01)	9.4 (1.13)
	95 % CI	(10.0, 10.8)	(9.0, 10.4)	(8.7, 10.1)

Spring Harbor 2007

This grantee provided training and technical assistance to multiple primary care provider practices that were implementing integrated behavioral health. The practices were provided templates for voluntarily monitoring process (numbers of patients screened and their demographics) and outcomes data (PHQ-9, GAD-7, and some quality of life measures). Participating practices had a very difficult time incorporating the template into their practices. Thus, most did not collect much process or outcome data.

Only four of the practices provided spreadsheets of data containing information for patients seen during the first half of 2010 (no data available for the second half of the year). The information provided included fairly complete data on the demographics of the patients they screened. The four sites were: Mid Coast Primary Care (84 patients screened), Oxford Hills Family Practice (60 patients screened), Portland Community Health Center (CHC) (32 patients screened), and Portland Family Medicine (FM) (156 patients screened). Of the 332 patients screened, nearly all (326, or 98%) were adults and in the age range of 19-64 (92%). The gender breakdown was relatively consistent across sites, about 1/3 (32%) male to 2/3 female (68%).

The PHQ-9 data provided by these four practices were more limited. In terms of initial assessments, Oxford Hills only provided a summary value of “[screen] positive/negative” but not the actual score. Portland CHC and Portland FM also provided the summary value “[screen] positive/negative” and only the scores for positives. Mid Coast provided the summary value and the actual initial scores for nearly all of their patients. A positive PHQ-9 score was defined as 5 points or higher (mild symptoms or higher) for all sites.

The screen positive rate for depression symptoms varied statistically significantly as well as substantively across sites. This may have been due to differences in depression prevalence across sites and/or differences in screening practices across sites. There also may have been variation in who was given a PHQ-9 to complete, such as whether patients were selected based on having a history of, or having current, symptoms.

Select Spring Harbor Learning Collaborative Sites: Initial PHQ-9 Depression Scores

	# of Patients Screened with PHQ-9 (1st half, 2010)	Initial PHQ-9 screen Positive	Initial PHQ-9 Screen Negative
Mid Coast Primary Care	81	84% (68)	16% (13)
Oxford Hills	60	15% (9)	85% (51)
Portland CHC	32	69% (22)	31% (10)
Portland FM	150	45% (68)	55% (82)
TOTAL	323	52% (167)	48% (156)

Neither Portland FM nor Oxford Hills provided any follow-up data; Portland CHC provided follow-up scores for two patients (not reported). Mid Coast had follow-up data for 40% (34) of their 84 patients. The Mid Coast site provides an opportunity to look at PHQ-9 scores in more detail.

Mid Coast patients who had a follow-up PHQ-9 score had higher initial scores than those who did not have follow-up data. There was a non-statistically significant trend for a reduction of 2.7 points in the average amount of depression symptoms over the course of about one month. Eleven of the 34 patients (32%) had a greater than 50% reduction in depression symptoms, a common benchmark for improved health.

Mid Coast Primary Care PHQ-9 Depression Scores Over Time

Site	Statistic	Initial PHQ-9	Follow-up PHQ-9	Average # Days of Follow-up	Number (%) of Patients w/ Complete Data
Mid Coast Primary Care	Mean	13.8	11.1	31.6 days	34 of 84 (40%)
	Standard Deviation	5.6	6.6		
	95% CI of mean	(11.9, 15.8)	(8.7, 13.4)		

York County Community Health Care (York)

York County Community Health Care is a FQHC that primarily serves homeless people and people living in public housing, in Sanford, Maine. All new patients were screened with the PHQ-9 for depression, but follow-up data collected only on a subset of that group (likely, those in treatment). There were 246 patients with an initial PHQ-9 score in 2010 in York's outcome data set. The number of patients with an initial PHQ-9 score was consistent with process data collected in the CDE database, which indicated that 222 patients were screened (of whom 35 were referred for additional integrated services). The numbers may not match exactly because of inaccurate counts in either data set, but the numbers are close enough to consider the outcomes data representative of the screened group.

The average (standard deviation) PHQ-9 score of patients was 12.9 (9.4), indicative of moderate depression symptomatology. Overall, 63% (155 patients) had moderate, moderate-to-severe, or severe symptoms of depression.

York, 2010: Initial PHQ-9 Score Distribution

Depression Symptoms (PHQ-9 Score)	Number of Patients	Percent of Patients
None (0)	60	24.4%
Minimal (1-4)	7	2.8%
Mild (5-9)	24	9.8%
Moderate (10-14)	37	15.0%
Moderate to Severe (15-19)	43	17.5%
Severe (20-27)	75	30.5%
Total	246	

Average (standard deviation) PHQ-9 Score: 12.9 (9.4)

There were 26 patients who had a second PHQ-9 score, representing only 10.5% of patients with an initial score. The average amount of time between measurements was 111.5 days, ranging from 7 days to 241 days. On average, initial PHQ-9 scores for those who also had a follow-up assessment were higher (16.7) than the entire sample (12.9). The average follow-up PHQ-9 score was significantly lower (12.7) by 4 points. While encouraging for those patients, it is not possible to know if their results are generalizable to the entire group of screened (or referred) patients.

2008 Clinical Implementation Grantees

Acadia Hospital

Acadia Hospital, an inpatient mental health and substance abuse facility, created a primary care clinic to assist access for their patients with complex behavioral health needs. The target population for this project was outpatient services clients of Acadia Hospital with both complex physical and mental health needs. During 2010-2011, according to the CDE database, 169 clients were referred to the primary care service embedded at the hospital. The SF12 version 2 was the outcomes assessment instrument for the project, providing two measures of functional health status, the Physical Health Score (PHS) and the Mental Health Score (MHS). The baseline SF12v2 was completed during the initial primary care visit,

and a follow-up assessment was scheduled for six months hence. Client scores were calculated normed to the US population (average score of 50 with a range from 0, or worst, to 100, or best, functioning).

Most clients, 111 (66%) completed an initial SF12v2. Baseline values showed that participants were significantly less healthy than the general population both in terms of mental and physical health. The mean MHS was 31.9 points (95% confidence interval: 29.7 to 34.2) and the mean PHS was 40.5 points (95% confidence interval: 38.1 to 42.9) – significantly lower than the 50 point US-normed average.

Acadia Hospital Clients: Baseline Mental and Physical Functioning

Scale	# Clients	Median	Mean	95% CI	Minimum	Maximum
MHS	111	30	31.9	(29.7, 34.2)	5	70
PHS	110	39	40.5	(38.1, 42.9)	18	70

A subsequent SF12v2 assessment was completed an average of 174 days later (range 135 to 202) for 17 clients. The low follow-up rate was due to the high no-show rate for clinic visits. Project staff switched to a mail out/mail back survey with a \$5 Wal-Mart card incentive, but it was not successful in improving the response rate. The distribution of change scores on the PHS was variable, with 9 clients having substantial declines (more than 5 points), 1 having substantial improvement, and the remainder no change. The distribution of change scores on the MHS was also variable, though not as much as the PHS. Four clients had substantial declines, 7 had substantial improvements, and 6 had no change in MHS. It is not possible to draw any conclusion about the effect of the program from these data, since it is plausible that clients who felt more poorly were willing to complete and return the survey or keep their clinic appointments. It is reasonable to conclude that the population had significant functional impairment and that, at least for some, the level of impairment can change substantially over several months.

Downeast Health Services

This project offered early intervention and wrap-around services for high-risk infants and young children (0-5 years old) and their families through collaboration with a number of service providers in Washington County. Infant and family support specialists were integrated into three primary care centers and one substance abuse treatment center. These specialists were supervised by the director of the Community Caring Collaborative. Millbridge Medical Center and Indian Township (IT) Health Center (Indian Health Service, Passamaquoddy Tribe-Motahkomikuk) each participated from 2009-2011, Pleasant Point (PP) Health Center (Indian Health Service, Passamaquoddy Tribe-Sipayik) participated from 2010-2011, and Discovery House, a methadone clinic, participated in 2009. Outcomes data were collected for Millbridge Medical Center and Indian Township and Pleasant Point Health Centers. Outcomes data were provided for 94 clients, which is relatively consistent with reporting in the CDE database (85 clients assessed for services). Discrepancies were due to small amounts of misclassification of patients as being associated with IT or PP or vice versa.

A treatment plan was developed for each client, with up to three goals set in each plan. For clients with multiple goals, sometimes goals were set at the same time, and sometimes they were set sequentially.



The outcome measure chosen for this project was whether or not the goals were met or there was improvement toward meeting goals. Clients were followed for an average of 102 days; those with more goals were followed longer than those with only one goal.

Over one-third of clients attained all of their goals (38%, 36/94), and 75% of them improved upon or attained at least one goal. The types of issues addressed in the treatment plan were many and varied, but included issues relating to child behavior (21%), case/care management (29%), family dynamics/home environment (16%), prenatal health or child birth (4%), or general mental or physical health issues (30%).

Meeting Goals Set in Treatment Plans (94 Clients)

# Goals Set	# Clients	Avg. # Days Follow-up	# (%) Attained All Goals	# (%) Improved on or Attained >=1 Goal
One	44	58	20 (45%)	23 (52%)
Two	30	140	12 (40%)	30 (100%)
Three	20	136	4 (20%)	20 (100%)
Overall	94	102	36 (38%)	73 (78%)

Type of Issues Addressed in Treatment Plans (164 Issues listed)

Issue Category	Examples	# (%) of Issues
Child Behavior	Behavior concerns, tantrums, improve school performance	35 (21%)
Case/Care Management	Get assessment, transportation, arrange for counseling	47 (29%)
Family Dynamics/Home Environment	Family dynamics, domestic violence, encourage safe environment, positive reinforcement	26 (16%)
Prenatal Health/Childbirth	Healthy birth, go to 6 week checkup	7 (4%)
Mental or Physical Health	Autism Screen, improve self-esteem, anxiety	49 (30%)

Franklin Health Pediatrics (FHP)

One of the goals of this project was to screen 75% of all children in Greater Franklin County for autism. This goal was to be accomplished by working with participating pediatric and family medicine practices to make the Modified Checklist for Autism-Toddlers (MCHAT) screening test part of the 18- and/or 24-month well child visit. The MCHAT was completed by parents and scored by clinical staff. FHP incorporated the MCHAT into their practice in 2010. In 2011, the practice screening rate, if not the county rate, reached the 75% mark. Because FHP could not break out well-child visits from acute visits, the screening rate may have been higher. Over the two-year period, the screening process also generated a total of 18 referrals (as reported in the CDE as well as the outcomes report) from the practices to Franklin Health Child and Adolescent Developmental Pediatrics for further autism assessment and services. Three other, smaller, FHP practices in Farmington, Livermore, and Wilton also began participating in the project in 2011. Only the Farmington Practice was able to provide screening

data; during the second half of 2010, 15 of 24 children (62.5%) aged 18-24 months who had a visit were screened using the MCHAT.

Franklin Health Pediatrics Autism Screening

	2010			2011		
	<i>Jan-June</i>	<i>July-Dec</i>	<i>Total</i>	<i>Jan-June</i>	<i>July-Dec</i>	<i>Total</i>
# Toddlers Screened with MCHAT	67	161	228	228	199	427
# Toddlers Seen at Clinic	N/A	259	N/A	262	298	560
Approx. Screening Rate	N/A	62%	N/A	87%	67%	76%
# Referrals to FH Child & Adolescent Developmental Pediatrics	2	4	6	8	4	12

Northeast Integrated Geriatrics Care

The goal of this project was to better address the behavior health needs of older adults in the acute hospital setting at Eastern Maine Medical Center and at four nursing homes that are part of Rosscare Nursing Homes, Inc. A new unit called Westside Court (WC) was designed at EMMC and staffed with clinicians trained to work with older adults with acute BH needs. An evaluation by the grantee of the hospital setting found a marked reduction in average length of stay from 45 days to 8.6 days² in 2010, and down to 6 days by the end of 2011. Also no anti-psychotic medications were used “as needed” (PRN) during the period of January 2010 through June 2011³.

During 2010-2011, the grantee also collected client-level data for residents of the four participating nursing homes. A LCSW divided her time among the four homes, and provided on-site counseling services to residents and support to staff on behavioral issues. Centers for Medicare and Medicaid Services’ (CMS) Minimum Data Set (MDS) is a required, standardized assessment for nursing home residents and is used for evaluating quality of care. In 2011, CMS revised the MDS very substantially (version 3.0), intending to make the assessment more reliable and accurate and inclusive of the patient’s voice. Version 2.0 of the MDS was based solely on observation, whereas version 3.0 was also based on interview. The project outcome measure in 2010 was one item from MDS 2.0, an observation that the resident had “a sad or pained facial expression,” a proxy for depression. In 2011, the outcome was one item from MDS 3.0, the resident response to a yes/no question about feeling down, depressed, or hopeless (and if yes, the number of days over the past two weeks feeling this way). Depression was not the only reason for referral to the LCSW, so not everyone was initially positive on these MDS measures.

² Comparing the length of stay of eligible patients in prior years to patients admitted to WC during project year.

³ For more details see the Case Study written about this grantee on the MeHAF website:

<http://www.mehaf.org/blog/2012/03/23/how-integrative-care-improves-life-nursing-home-patients-mental-health-diagnoses/>

Data were provided for 74 residents, nearly evenly split between the two reporting years. The number of residents was relatively consistent with the counts of residents reported in the CDE database (88). Follow-up data was available for 61 (82%) of these residents; most of the other 13 residents were discharged before follow-up assessment occurred. The average time between baseline and follow-up assessment was 36 days (ranging from 4 to 92 days).

Number of Residents Who Saw the LCSW and was Assessed Using the MDS

Nursing Home	# Residents in 2010	# Residents in 2011	Total # residents	Total # of residents with 2 assessments	Average # days between assessments
Colonial	8	10	18	13	50
Dexter	3	4	7	7	35
Ross Manor	20	10	30	26	33
Stillwater	7	12	19	15	28
Total	38	36	74	61	36

Given the relatively small number of cases per home and since the same LCSW implemented the same program in all four locations, results are presented for all homes combined. In 2010, among 30 residents with two assessments, 12 (40%) were observed with sad expression at baseline and 11 were similarly observed at follow-up. Two residents' expressions got better, another became sad, and the rest stayed the same. In 2011, 52% (16/31) of residents responded they felt down/depressed or hopeless on the baseline assessment. At follow-up, that number had fallen to 35% (11/31); 5 residents no longer reported this symptom, 3 had the symptom but for fewer days; there were no new cases of feeling down/depressed/hopeless among the LCSW caseload. The seemingly stronger results in 2011 are likely due to the improved measurement system, better able to reflect the impact of the intervention.

Depression-related Outcome Measures for Nursing Home Residents

	All Patients	Only Patients with Two Assessments	
	Initial Assessment	Initial Assessment	Follow-up Assessment
MDS 2.0 (2010) # (%) sad, pained facial expression observed	17/38 (45%)	12/30 (40%)	11/30 (36%)
MDS 3.0 (2011) # (%) answered yes to feeling down/depressed/hopeless	18/36 (50%)	16/31 (52%)	11/30 (36%)
<i>Of those who answered yes, # days felt this way during the past two weeks:</i>			
2-6 days	9/18 (50%)	7/16 (44%)	7/11 (63%)
7-11 days	4/18 (22%)	4/16 (25%)	2/11 (18%)
12-14 days	5/18 (28%)	5/16 (31%)	2/11 (18%)

Spring Harbor 2008

This integration project was a consultation model, adapting the Massachusetts Child Psychiatry Access Project (CPAP). The model was adapted in that the CPAP team was comprised of a psychiatrist and clinical care coordinator but not a social worker as in the MA model. This team provided consultation services to four pediatric practices serving patients in Brunswick and Topsham (Martin's Point, Bowdoin Medical, Parkview Pediatrics, and Mid Coast Pediatrics). Pediatricians could call the CPAP and get a call back within an hour from the clinical care coordinator, who would help specify the issue and determine the best approach to addressing it. The approaches included a telephone consultation between the PCP and the child psychiatrist, the clinical care coordinator working directly with the family to access local behavioral health services, and referral to the CPAP psychiatrist for a face-to-face evaluation of the patient (either for ongoing care by the PCP or by a local mental health provider).

In order to assess the quality of the consultation, pediatricians were surveyed twice, in June of 2010 and in January of 2011. Fifteen clinicians from the four practices completed the survey in either period; almost half (7/15) completed it both times. Most clinicians (80% or 12 out of 15) used CPAP during the six months prior to completing a survey.

All 12 pediatricians who used CPAP agreed that they found the consultation helpful as well as timely. Among the seven clinicians who took the survey twice, 5 increased their ratings for helpfulness to "strongly agree," and 4 increased their ratings for timeliness to "strongly agree."

In a seeming contradiction, scores on agreeing with the statement of there being "adequate access to child psychiatry" were lower in January 2011, than in June 2010, among pediatricians using CPAP. The mean score decreased from 4.0 to 2.7 on the 5-point scale that ranged from strongly disagree (1) to strongly agree (5). Information from the four clinicians from Bowdoin Medical and 2 of 4 clinicians from Martin's Point who responded to both surveys and provided comments help explain these results. It seems that as clinicians became more experienced using the program, they wanted greater access to it. One respondent noted: "There is a wait list now!" Another commented: "Still not enough child services – psychiatry and counselors – for less than 8 year olds. CPAP is a great program. It would be great to have someone on site monthly to help with evaluations. Scheduling in primary care makes it hard to address psychiatric problems. With phone back up, I feel more confident with difficult patients."

The second page of the survey ascertained how a pediatrician's own practice was impacted by CPAP. More methods of behavioral health screening were reported in January 2011 than in June 2010 (increasing, on average from 1 to 3 methods). The number of pediatricians using formal screening tools increased from 4 in June to 11 in January. There were 72 telephone consultations during the six months prior to the June survey compared to 24 consults prior to the January survey. There were 11 face-to-face consultations during both time periods. More screening and less consultation are consistent with the idea that the CPAP resource may have become more constrained over the course of the project.

Most of the pediatricians who used CPAP reported following up with all or most of the patients they had referred, seeing them for the BH/MH issue after the consultation (67%, or 8/12). Four pediatricians (33%) did not answer these questions.

Tri-County Mental Health Services (TCMHS)

In 2008 this project provided behavioral health services in three primary care practices through a contractual relationship between Swift River Health Care (SRHC) and TCMHS. TCMHS provided the services of two licensed clinicians to SRHC and streamlined referrals to other TCMHS services in the community. The participating family care practices were Swift River Health Care, Dixfield Health Center, and River Valley Internal Medicine. This program focused on patients with high mental health needs and low physical health needs. The outcome measures chosen for the project were PHQ-9 for depression and the GAD-7 for anxiety for adults and the Pediatric Symptom Checklist (PSC) for children.

Over the course of 2010-2011, outcomes data (either initial or follow-up scores or both) were collected on 345 adults, consistent with reporting in the CDE database through the second quarter of 2011 (not all three practices supplied outcomes data for the second half of 2011). Among these adults 264 had initial depression and anxiety scores. Baseline symptoms tended to be moderate to severe; 65% (171/264) of patients had both moderate to severe depression and anxiety. A similar pattern was seen among adults assessed through the Pen Bay integration project, noted earlier in the report.

Initial Depression and Anxiety Levels of TCMHS/SRHC Patients (n=264)

	No/Mild Depression (GAD-7 0-9)	Moderate/Severe Depression (GAD-7 >9)
No/Mild Depression (PHQ-9 0-9)	52 (20%)	23 (9%)
Moderate/Severe Depression (PHQ-9 > 9)	17 (7%)	171 (65%)

Follow-up assessments were often not collected, in part because some patients stopped accessing services after only a few sessions. Only 64 patients (out of 264, or 24%) also had follow-up scores, which were measured, on average, 97 days after the initial assessment. Results for these patients was encouraging; the average change in PHQ-9 score was a reduction of nearly 6 points (-5.7 change; 95% confidence interval: -6.9 to -4.5 points) and the average change in GAD-7 was a reduction of 4 points (-4.0 change; 95% confidence interval: -5.0 to -3.0 points). While the baseline scores of these were similar to those who did not have follow-up assessment, it is not possible to know if this subset of patients is representative of all patients.

TCMHS/SRHC PHQ-9 and GAD-7 Change Scores

	Scale	Mean	95% CI	Minimum	Maximum
PHQ-9 for Depression	Baseline	13.6	(12.2, 15.0)	1.0	25.0
	Follow-up	8.0	(6.9, 9.2)	0.0	17.0
	Change	- 5.7	(-6.9, -4.5)	-20.0	6.0
GAD-7 for Anxiety	Baseline	11.3	(10.1, 12.5)	2.0	21.0
	Follow-up	7.4	(6.3, 8.5)	0.0	21.0
	Change	- 4.0	(-5.0, -3.0)	-12.0	11.0

Bridgton Internal Medicine

TCMHS also received a 2009 implementation grant, and created another integrated care partnership with Bridgton Internal Medicine. The same care model and outcomes measures were implemented at this site and the results were similar as well, indicating moderate-to-severe baseline depression and anxiety symptoms, with a high proportion of people with co-occurring conditions.

Over the course of 2010-2012, outcomes data (either initial baseline scores or both baseline and follow-up scores) were collected on 435 adults. Of the 424 with reliable data, all had an initial PHQ-9 assessment only and 355 had both an initial PHQ-9 and GAD-7 assessment. The table below shows the baseline scores for the two outcome measures separately.

PHQ-9 and GAD-7 Baseline Scores of Bridgton Patients

	Total Baseline Scores	Moderate and Severe Scores	Mean	95% CI	Minimum	Maximum
PHQ-9 for Depression	424	310 (73%)	13.3	(12.8, 13.9)	0.0	27.0
GAD-7 for Anxiety	355	216 (61%)	11.2	(10.6, 11.7)	0.0	21.0

For the 355 adults who had *both* initial depression and anxiety scores, 50% (179/355) of patients had both moderate to severe depression and anxiety at baseline.

Initial Depression and Anxiety Levels of Bridgton Patients at Baseline (n=355)

	No/Mild Depression (GAD-7 0-9)	Moderate/Severe Depression (GAD-7 >9)
No/Mild Depression (PHQ-9 0-9)	69 (19%)	37 (10%)
Moderate/Severe Depression (PHQ-9 > 9)	70 (20%)	179 (50%)

Only 150 patients (out of 424, or 35%) also had follow-up PHQ-9 scores, measured an average of 55.7 days after the initial assessment. Results for these patients were encouraging; the average change in PHQ-9 score was a reduction of over 4 points (-4.2 change; 95% confidence interval: -4.9 to -3.5 points). For the GAD-7, 108 patients (out of 355, or 30%) also had follow-up scores, which were measured, on average 54.5 days after the initial assessment. Like the PHQ-9, the average change in GAD-7 scores was promising, with a symptom reduction of over 3 points (-3.3 change; 95% confidence interval: -4.1 to -2.5 points). These improvements in symptoms were similar to those seen among patients from the other TCMHS sites, if slightly lower. On average, the time to the follow-up visit was shorter at Bridgton (mean 55-56 days) compared to the other TCMHS sites (mean 97 days). Shorter duration of follow-up allows for less time for improvement, particularly for more serious cases.

2009 Clinical Implementation Grantees

Eastern Maine Medical Center (Center for Family Medicine)

The Center for Family Medicine (CFM) serves as a residency training program as part of Eastern Maine Medical Center (EMMC). The purpose of the MeHAF grant was to incorporate a screening, brief intervention, referral, and treatment (SBIRT) algorithm into its electronic medical record and clinical practice. CFM had two behavioral health providers who assisted with training and precepting residents and who were an integral part of the care team. Another key component of the original project was developing a “wellness survey” that incorporated 3 alcohol-use and 3 drug-use screening questions, which in turn was to be implemented in on-line format in a kiosk in the waiting room. The data from the wellness survey was to populate the medical record so the provider could incorporate further evaluation or intervention as necessary. The target population was CFM patients coming in for their annual physical. Follow-up assessment was planned to occur when the patient came back a year later for a physical and again completed the wellness survey’s alcohol and drug use screening questions.

There were technical complications with regards to setting up the kiosk with the survey as well as linking the survey data to the electronic medical record (Centricity). The start date for engaging patients in the project was pushed back significantly into the second half of 2011 as attempts were made to resolve the technical complications. Thus, few patients who were initially screened in 2011 were eligible for annual follow-up screening during the second half of 2012 (the last data collection period of the evaluation).

The grantee was able to collect process measures for the CDE data base, tracking the number of people screened, assessed, and referred for services from October 2011 through December 2012. More detailed process data was collected and submitted with the CDE data. That information is presented here, providing a unique look at substance use/abuse screening.

Only 82 of 7,077 patients were screened using the wellness survey, whereas far more of them were screened by providers: 2,841 for alcohol use and 3,994 for drug use. The wellness survey identified 22.0% (n=18) patients at risk for drug or alcohol misuse/abuse. Providers identified 30.6% (n=869) patients for alcohol misuse/abuse and 5.6% (n=222) for drug abuse. According to SAMHSA’s National Survey on Drug Use and Health (NSDUH 2010), an estimated 8.9% of Americans 12 years or older used an illicit drug or abused a psychotherapeutic medication⁴. Also, an estimated 23.1% of Americans aged 12 years or older engaged in binge drinking and 6.7% engaged in heavy drinking, and among 16.9 million heavy drinkers aged 18+, 31.8% were illicit drug users⁵. Using the NSDUH statistics as a rough guide, if we assume a 1/3 overlap between at-risk drug and alcohol use, then the combined screened positive rate for the physicians would be about 24%, qualitatively similar to the rate from the wellness survey.

⁴ <http://www.samhsa.gov/data/NSDUH/2k10ResultsRev/NSDUHresultsRev2010.pdf>

⁵ <http://www.samhsa.gov/data/nsduh/2k10nsduh/2k10results.htm#3.1.9>

EMMC Screening Results for Substance Use, by Method of Screening

		Screened via Wellness Survey (WS)			Screened via Provider (PCP)					
Time Period	Total Visits for Quarter	# Screened (for Alcohol or Drug Use)	# Positive	% Positive	# Screened for alcohol use	# Alcohol Positive	% Alcohol Positive	# Screened for Drug Use	# Drug Use Positive	% Drug Use Positive
4 th Quarter, 2011	857	23	6	26.1%	278	101	36.3%	138	10	7.2%
1 st Quarter, 2012	1436	16	2	12.5%	648	202	31.2%	710	31	4.4%
2 nd Quarter, 2012	1874	12	5	41.7%	754	204	27.1%	1145	65	5.7%
3 rd Quarter, 2012	1426	17	5	29.4%	568	172	30.3%	1015	64	6.3%
4 th Quarter, 2012	1484	14	0	0.0%	593	190	32.0%	986	52	5.3%
Total Screened	7077	82	18	22.0%	2841	869	30.6%	3994	222	5.6%

However, the likelihood of screening positively and being referred to BH for assessment was substantively higher for PCP-screened patients (10.7%, 117/1091; 1091=869+222) than for those screened positive via wellness survey (5.6%, 1/18). This is somewhat of an underestimate for PCP-screened patients, because we were unable to account for the overlap among those who were positive for both alcohol and drugs in the denominator.

Referrals to Behavioral Health Providers

Time Period	# Screened Positive (via WS) and Referred to BH for Assessment	% Screened Positive (via WS) and Referred to BH for Assessment	# Screened Positive (via PCP) and Referred to BH for Assessment	% Screened Positive (via PCP) and Referred to BH for Assessment	Total # of Referrals to BH	# Patients who did not show up for Assessment	# Kept Assessment Appointment, no further treatment needed	# Kept Assessment Appointment and referred for more treatment
4 th Quarter, 2011	0	0.0%	21	19.9%	21	0	14	7
1 st Quarter, 2012	0	0.0%	15	6.4%	15	1	12	2
2 nd Quarter, 2012	1	20.0%	28	10.4%	29	2	4	23
3 rd Quarter, 2012	0	0.0%	25	10.6%	25	0	0	25
4 th Quarter, 2012	0	--	28	11.6%	28	0	0	28
Total Screened	1	5.6%	117	10.7%	118	3	30	85

CFM behavioral health clinicians also provided consultations to the resident physicians, advising them on how to talk with their patients about substance abuse issues. They provided 210 consultations, lasting an average of 11 minutes per consultation (total level of effort was 2,380 minutes or 40 hours) for patients screened positively/at risk for substance abuse. An additional 47 consultations were provided to the residents regarding patients who had a mental health diagnosis and who were at-risk for substance use.

Mercy (Fore River Clinic)

The Fore River Clinic is a Mercy Hospital-affiliated primary care practice in Portland. The goal of its project was to provide family-centered care to patients with dual diagnoses, defined as having both a chronic physical and mental health condition. The conditions were hypertension, low back pain, and diabetes co-occurring with depression and/or anxiety. As part of this project, a LCSW was hired to help assess and care for these complex patients. The clinic already employed a Somali Care Coordinator (funded by Maine's Department of Health and Human Services), and a second patient care navigator was hired from Amistad to assist other clinic patients as well.

During 2010-2012, 28 patients were enrolled in the hypertension group (HTN), 32 in the low back pain group (LBP), and 34 in the diabetes group (DM). There were 11 native Somali- or Arabic- speaking patients enrolled (12% of the 94 total patients). The majority were seen several times over the course of over several months (up to a year or so). At nearly every visit, a condition-specific measurement (blood pressure, self-reported pain on a 10-point scale, or HbA1c, for the HTN, LBP, DM groups respectively) was recorded. Most of the time, the PHQ-9 depression scale and Beck Anxiety Inventory scores were also recorded. Some patients resisted completing the mental health scales, particularly if they were seen frequently in clinic or were of Somali origin. The Beck scale was the more problematic of the two and sometimes the GAD-7 was successfully substituted.

In order to describe briefly and clearly the outcomes for this project, patients who had been enrolled six months or longer were selected. At the time of their initial assessment, the majority of patients had moderate or greater levels of depression (57% of HTN group, 65% of LBP group, 46% of DM group and high anxiety levels. One-third of HTN patients (9/27) had blood pressure readings in the hypertension range; 62% of DB patients (16/26) had HbA1c levels of over 7%. The average pain score was 6.9 on a 10-point scale among LBP patients.

Not all of these selected patients had complete initial and follow-up data on condition-specific and mental health outcomes. Thus, the subgroup with complete data is described in the tables below. This subgroup exhibited greater initial depression and anxiety than the group as a whole. Over six months, the HTN and LBP groups each averaged modest improvement in their condition-specific measure. Fewer had blood pressure readings in the hypertension range (5/27=19%, versus 9/27 initially) and the pain score decreased to an average of 6.5 points (down from 6.9 points initially).

Mercy Fore River Clinic Dual Diagnosis Patients

Condition	MH Severity*	Baseline Depression	Baseline Anxiety	Condition Specific Outcomes
Hypertension	Minimal	1 (7%)	2 (13%)	<u>Mean Blood Pressure:</u> Initial: 126.5/78.1 Follow-up: 123.7/76.1 <u>% HTN:</u> Initial: 9/27 Follow-up: 5/27 (n=27)
	Mild	2 (13%)	0 (0%)	
	Moderate	3 (20%)	6 (40%)	
	Moderate-to-Severe	5 (33%)	N/A	
	Severe	4 (27%)	7 (46%)	
Low Back Pain	Minimal	0 (0%)	1 (6%)	<u>Mean Pain Level (n=20):</u> Initial: 6.9 Followup: 6.5
	Mild	5 (29%)	4 (24%)	
	Moderate	1 (6%)	3 (18%)	
	Moderate-to-Severe	3 (11%)	N/A	
	Severe	8 (47%)	9 (53%)	
Diabetes	Minimal	1 (4%)	2 (15%)	<u>Mean HbA1c:</u> Initial: 8.6% Follow-up: 8.0% <u>% HbA1c>7%:</u> Initial: 9/26 Follow-up: 8/26 (n=26)
	Mild	2 (15%)	3 (23%)	
	Moderate	4 (31%)	3 (23%)	
	Moderate-to-Severe	4 (31%)	N/A	
	Severe	2 (15%)	5 (38%)	

* MH results are presented for patients with complete MH and condition-specific data, a smaller number than those who had condition-specific data. These sub-sets had similar initial condition-specific results to their corresponding full group results.

A unique aspect of this project is the ability to describe the correlation between condition-specific symptoms and depression⁶ (PHQ-9) on 45 patients with complete data (15 in HTN group, 17 in LBP group, and 13 in DM group). There was no significant correlation between initial PHQ-9 depression scores and initial systolic blood pressure, diastolic blood pressure, or blood glucose (HbA1c) readings. However, there was a significant positive correlation between pain scores and depression in the LBP group (spearman correlation⁷: $r_s = 0.60$, $p\text{-value} = 0.01$). A positive correlation means those with greater (lower) pain tended to have greater (lower) depression levels. At follow-up, the correlation was no longer significant ($r_s = 0.34$, $p\text{-value} = 0.18$).

To try to understand the correlation pattern for the low back pain group, we examined change in pain, depression, and anxiety symptoms over the approximately 6-month follow-up period. The 17 LBP patients with complete data were grouped according to whether their follow-up pain rating was more

⁶ Because a mix of GAD-7 and Beck scores were reported for anxiety, correlation values could not be computed between condition measures and anxiety. Although the actual GAD-7 and Beck scores are not directly comparable, patients could be classified into similar anxiety symptom categories using either scale (see tables).

⁷ Spearman correlation coefficient is the small-sample analog to the more commonly known Pearson product-moment correlation. Correlation scores can run from -1 to +1. A zero correlation coefficient means there is no association between the two variables.

than a point lower (less pain), within a point (same pain), or more than a point higher (more pain) than initially. Within each of these groups, changes in depression and anxiety symptoms were noted. Having greater (less) depression was defined as being classified in a higher (lower) depression symptom category at follow-up than initially. Having the same depression severity was defined as being classified in the same symptom category at both points in time. The same type of categorization was performed for anxiety symptoms as well.

Just over one-third (6/17) were experiencing less pain, 41% the same pain (7/17), and 24% (4/17) more pain over six months. Pain was still substantial, on average, in all three groups. Among those with less pain, 4/6 also had less depression and anxiety. Among those with the same pain, 3/7 had less depression (2 had also had less anxiety). Among those with more pain, 2/4 had either less anxiety or less depression symptoms. The overall pattern in this small group is that several patients had less mental health symptoms even if pain had not abated, perhaps indicative of improved resilience or coping.

Low Back Pain Patients: Change in Pain and Mental Health (n=17)

Pain Change Category	Mean Pain Level		Change in Depression/Anxiety Symptoms	
	Initial	Follow-up		
Less Pain (6, 35%)	7.7	5.0	Less/Less	3
			Same/Less	1
			Same/Same	1
			More/More	1
Same Pain (7, 41%)	7.1	7.4	Less/Less	2
			Less/Same	1
			Same/Same	2
			More/Less	1
			More/More	1
More Pain (4, 24%)	5.0	8.4	Same/Less	1
			Less/Same	1
			Less/More	1
			More/More	1

This analysis relates back to the correlation statistics between pain and depression mentioned above. At baseline, pain level and depression level were clearly correlated, implying greater depression symptoms among those with greater pain levels. However at follow-up, some patients had lower depression scores for a given level of pain (particularly among those with similar or worse pain), which weakened the correlation.

Conclusion

As of December, 2012, approximately 11,124 patients had been assessed for integrated primary care - behavioral/mental health services among 24 implementation grantees. Results summarized in this report begin to address the question: “What was the impact of integrated services on patients’ health?” There are many ways to address this question. The one chosen for this initiative was to try to assess changes in patient-level health status over a 3- to 6- month timeframe, for clients who were assessed

and referred for integrated services (typically behavioral/mental health services or consultation embedded in primary care).

Not all grantees were able to collect the initial and follow-up assessment data necessary to estimate change in health status. Almost half (46%, 11/24) of grantees were able to provide initial and follow-up assessments on at least some of their patients and thus describe change in health status. About one-third (38%, 9/24) were only able to provide initial assessments or screening data and, thus, describe baseline severity. The remaining 4 grantees (17%) were not able to provide any data for this evaluation. Thus, there is no one summary cross-site estimate of the impact of integrated services on health status. However, given the intentionally varied nature of these programs and that they were geared toward maximal implementation rather than research, this was expected. Insights about the health status of patients can still be gleaned from the data, as well as about the factors that promote and inhibit outcomes data collection.

Based on initial assessment data from grantees monitoring mental health, patients had substantial symptoms. Five grantees used the PHQ-9 scale to assess depression and all groups averaged scores in the 12-14 point range, indicative of moderately severe symptoms. The table below differs a bit from the grantee-specific analysis in that it describes on all patients with an initial PHQ-9 assessment, whereas some of the grantee specific analyses focused on the subset of these patients who also had either: (1) a follow-up PHQ-9 score; or (2) had an initial anxiety assessment.

Cross-Grantee Average PHQ-9 Scores: Assessment of Initial Severity of Depression

Grantee	# Patients	Average PHQ-9 Score
DFD Russell Medical Centers	133	13.7
Pen Bay Health Care	313	13.1
Spring Harbor 2007 (Mid Coast Primary Care)	84	11.9
York County Community Health Care	246	12.9
Tri-County Mental Health Services:		
- Swift River, Elsmore-Dixfield, River Valley (2008 grant)	264	13.6
- Bridgton (2009 Grant)	424	13.3
Mercy (Fore River Clinic)	69	13.3
Overall	1,533	13.2

Moderate-to-severe anxiety co-occurring with moderate-to-severe depression was also found among most patients enrolled through the Tri-County (2008 grant) (65%), Pen Bay (53%), and Mercy Fore River (67%) programs. Other grantees using different measures also found substantial initial functional impairment (Acadia 2008 and AMHC/Fish River Practice).

Over the course of 2011, residents at EMMC’s Center for Family Medicine screened nearly 4,000 patients for drug use and over 2,800 for alcohol use. They found substantial numbers who were at risk for substance use issues (30.6% for alcohol and 5.6% for other drugs); patients who could benefit from

further assessment with a behavioral health provider. Substance use was not the focus of other integration projects; these results imply other similar integrated practices could have similar prevalence.

Grantees whose practices were able to provide good follow-up data for all, or a well-defined segment, of their patients generally showed positive impacts on health status. DFD Russell's patients with major depression (133/167 with complete data) had a statistically-significant drop in symptoms (5.6 PHQ-9 points), with 48% achieving a 50% or more reduction in symptoms or a score of 5 or less (remission). Sacopee Valley Health Center's diabetes patients had a non-statistically significant drop in HbA1c among participants (16 who completed the program and 10 who transitioned out) from 10.4% to 9.2%, although none achieved a reading of less than 7% over 8-12 months. Downeast counselors worked with high-risk families with young children to develop treatment plans for counseling and other services. Over one-third of clients attained all of their goals (38%, 36/94) and 75% of them improved upon or attained at least one goal.

Results for on-site counseling at Rosscare Nursing Homes were clearer in 2011 than in 2010 after CMS's upgrade from the MDS 2.0 to the MDS 3.0 quality management data system. In 2011, 52% (16/ 31) residents responded that they felt down, depressed or hopeless, improving to 36% (11/30) at follow-up; among those with symptoms, fewer depressed days were reported (an estimated 8.1 days vs. 6.5 days). Thirty Amistad members were recruited to monitor their weight over a period of six months. Over the course of six months, 50% of clients had a substantial weight change in the preferred direction, either gaining more than 5 pounds (2 of 4 clients) or losing more than 5 pounds (9 of 18 clients). At Mercy's Fore River Clinic, a cohort of 94 patients with complex physical and mental health issues was established, including 11 patients who were immigrants from Somalia. With the majority of these patients, they were able to establish routine clinic visits to monitor their conditions and assess and address social service needs. Modest improvements were noted in health status, particularly in the mental health of low back pain patients.

Three grantees provided non-patient level outcomes information. Spring Harbor's 2008 Child Psychiatry Access Project surveyed participating pediatricians in four practices about the quality of the new consultation service. All 12 pediatricians who used CPAP agreed the psychiatry consultations were helpful and timely. The number of pediatricians using formal mental health screening tools increased from four to eleven. Two-thirds of clinicians followed up with their patients post-consultation. Community Dental successfully opened a new practice in Rumford. The common dental-physical-mental health referral form developed with CD community partners was discontinued, but Rumford Hospital's Emergency Department continued to be a significant source of dental referrals. Franklin Developmental Pediatrics worked with the Franklin Health practice in Farmington to implement the MCHAT autism screening tool during well-child visits; in 2011, at least 76% of children aged 18-24 months were screened. Eighteen referrals for autism were also made to Franklin Child and Adolescent Development Practice.

There were two factors that enabled grantees to report useful outcomes data. The first was having organizational experience with quality measurement and a project leader who worked with staff to plan

and implement outcomes assessment. Grantees with these capabilities were able to provide information even though none of them had a “point and click” reporting software application.

A second factor was the ability to track all assessed patients over time, or the ability to define a subset of patients and track them over time. Being able to define a group of patients greatly improves the ability to interpret outcomes. For example, having initial and follow-up assessments on only a small subset of patients begs the question of why that group had follow-up data. They could have been more compliant, more seriously ill or some other characteristic that makes them less generalizable to all patients. If on the other hand, initial and follow-up assessments were gathered on all patients with major depression (or diabetes or other well-defined characteristic), then interpretation is relatively more straightforward. The limitation is that some patient sub-groups would not be represented in the data. But with time, budget, and staff constraints, it is a reasonable approach, and consistent with a quality management framework. Over time different sub-groups could be assessed.

There were several factors that were particular barriers to reporting useful outcomes data. Patients who were very difficult to engage in care were also very difficult to engage in outcomes assessment. Even though some had strong assessment plans in place (Acadia, Penobscot, York), they served populations who were homeless and living with serious mental illness who often are more challenging to engage in care. A related issue is that it is difficult to establish an “appropriate” timeframe for health status improvement - and what improvement might mean - in health status for very complex patients, with multiple serious conditions (such as at Mercy’s Fore River Clinic). Finally, grantees whose clinical practice partners did not have the resources or other incentives to provide data did not do so (Spring Harbor 2007 and AMHC).