State of Colorado

Department of Health Care Policy & Financing

FY 06–07 PIP VALIDATION REPORT

Improving Outcomes for High-Risk Youth Through AFFIRM Care Management

for

Access Behavioral Care

June 2007

HSAG HEALTH SERVICES ADVISORY GROUP

1600 East Northern Avenue, Suite 100 • Phoenix, AZ 85020

Phone 602.264.6382 • Fax 602.241.0757

HSAG HEALTH SERVICES ADVISORY GROUP

CONTENTS

for Access Behavioral Care

Overview

The Balanced Budget Act (BBA) of 1997 (Public Law 105-33) requires that states conduct an annual evaluation of their managed care organizations (MCOs) and prepaid inpatient health plans (PIHPs) to determine the MCOs' and PIHPs' compliance with federal regulations and quality improvement standards. According to the BBA, the quality of health care delivered to Medicaid consumers in MCOs and PIHPs must be tracked, analyzed, and reported annually. The Colorado Department of Health Care Policy & Financing (the Department) has contractual requirements with each MCO and behavioral health organization (BHO) to conduct and submit performance improvement projects (PIPs) annually. As one of the mandatory external quality review activities under the BBA, the Department is required to validate the PIPs. To meet this validation requirement, the Department contracted with Health Services Advisory Group, Inc. (HSAG) as an external quality review organization. The primary objective of the PIP validation is to determine the compliance with requirements set forth in 42 CFR 438.240(b)(1), including:

- Measurement of performance using objective quality indicators.
- Implementation of system interventions to achieve improvement in quality.
- Evaluation of the effectiveness of the interventions.
- Planning and initiation of activities for increasing or sustaining improvement.

The Centers for Medicare & Medicaid Services (CMS) publication, *Validating Performance Improvement Projects: A Protocol for Use in Conducting Medicaid External Quality Review Activities*, Final Protocol, Version 1.0, May 1, 2002, was used in the evaluation and validation of the PIPs.

Summary of Study

The purpose of the study was to assess whether participation in a structured care management program could reduce psychiatric hospital readmissions and improve clinical functional outcomes for high-risk child and adolescent consumers.

Study Topic

The study topic addressed CMS' requirements related to quality of care and services. The topic addressed improving outcomes for high-risk child and adolescent consumers who received care through the use of the Access Family-Focused Intervention and Recovery Model (AFFIRM) Care Management Program.

Study Methodology

Two study indicators were developed to collect data that would answer the study question. Data were collected on psychiatric inpatient readmission rates at 30 days and on reduction of three- or six-month Child and Adolescent Needs and Strengths–Mental Health (CANS-MH) domain scores from initial CANS-MH domain scores. Administrative and manual data were collected to measure the outcomes. The study population for the first indicator consisted of **Access Behavioral Care** (**ABC**) youths 0 to 17 years of age who met study inclusion criteria. The population for the second indicator was a subset of AFFIRM program participants who met the inclusion criteria, were continuously enrolled, and actively participating in the AFFIRM program for at least three months, and had at least two CANS-MH assessments completed. A care management program aimed at providing enhanced clinical and social support to child and adolescent consumers was put into place to reduce 30-day readmission rates and CANS-MH scores. Data were collected and analyzed annually, and the results were used to demonstrate quality of care and services provided.

Study Results

For the FY 06–07 submission, **ABC** had collected three measurements of Study Indicator 1 and two measurements of Study Indicator 2. For the first indicator—psychiatric inpatient readmission rates at 30 days—there was statistically significant improvement from baseline to the second remeasurement. For the second indicator—reduction of three- or six-month CANS-MH domain scores from the initial CANS-MH domain scores—there was a substantial decline; however, the decline was not statistically significant. For the current measurement period, one of five, or 20 percent of consumers, showed a reduction of at least one point in four of six domains. For this assessment tool, a lower score indicates better performance.

Scoring

HSAG validates a total of 10 activities for each PIP. The PIP is validated annually. The validation reflects activities that have been completed. A health plan (BHO) may take up to three years to complete all 10 activities. Each activity consists of elements necessary for the successful completion of a valid PIP. Evaluation elements are the key CMS protocol components for each activity that reflect the intent of what is being measured and evaluated. Some of the elements are critical elements and must be scored as *Met* to produce an accurate and reliable PIP. Given the importance of critical elements, any critical element that receives a *Not Met* score results in an overall PIP validation status of *Not Met*. If one or more critical elements are *Partially Met*, but none is *Not Met*, the PIP will be considered valid with low confidence. Revisions and resubmission of the PIP would be required.

Summary of Validation Findings

- For this review, 10 activities with a total of 53 elements were validated. Of this number:
  - 42 evaluation elements were *Met*.
  - 2 evaluation elements were *Partially Met.*
  - 0 evaluation elements were *Not Met*.
  - 9 evaluation elements were *Not Applicable (N/A).*
- The total number of critical elements that were evaluated equaled 11. Of this number:
  - 9 critical elements were *Met*.
  - 0 critical elements were *Partially Met*
  - 0 critical elements were *Not Met*.
  - 2 critical elements were *N/A*.

The final validation finding for **ABC**'s PIP showed an overall score of 95 percent, a critical element score of 100 percent, and a *Met* validation status.

Conclusions

For this validation cycle, the study successfully addressed the quality of care and services delivered to consumers. The study topic was applicable to the **ABC** population; it reflected a high-risk, high-volume condition; and it had the potential to improve consumer health status and satisfaction. The study question set and maintained the focus of the study and was answerable using well-developed study indicators. There was statistically significant improvement in the rates of psychiatric inpatient readmissions at 30 days; however, **ABC** determined that the difference was clinically negligible due to the small number of consumers enrolled in the AFFIRM program. There was a substantial decline in three- or six-month CANS-MH domain scores from the initial CANS-MH domain scores; however, the small number of consumers enrolled for the measurement period prevented any conclusive judgments about the success of the interventions. As a result of deficiencies in the design of the AFFIRM intervention and the outcomes measured, **ABC** management determined it was no longer possible to continue the program as originally designed. **ABC** will develop a new PIP focused on a high-priority area that needs improvement in the quality of mental health care, and that is better aligned to meet the needs of its consumers.

Requirements

There were no requirements for this validation cycle.

Recommendations

HSAG recommends that **ABC** target improvement in relevant areas of clinical care. The topic should be identified through data collection and analysis of comprehensive consumer needs, care,

and services. The study topic should reflect the BHO's Medicaid enrollment in terms of demographic characteristics, prevalence of disease, and the potential consequences (risks) of the disease. The study topic should reflect high-volume or high-risk conditions, address a broad spectrum of care and services, and have the potential to affect consumer health, functional status, or satisfaction.

Comparison of Years 1 and 2

For the Year 1 validation cycle, Activities I, Appropriate Study Topic, through VII, Appropriate Improvement Strategies, were assessed because **ABC** had only completed intervention implementation at the time of the PIP submission. For Year 2, the PIP was assessed through Activity X, Sustained Improvement Achieved. **ABC** determined that although there was statistically significant improvement in readmission rates from baseline to the second remeasurement, the difference was clinically negligible. For Year 2, **ABC** had completed two measurements of Study Indicator 2. There was improvement in Study Indicator 2; however, the improvement was not statistically significant. There was a small number of consumers enrolled in the AFFIRM program for FY 06–07, and **ABC** determined that it would develop a new PIP that was better aligned with the needs of its consumers.

Validating PIPs involves a review of the following 10 activities:

- Activity I. Appropriate Study Topic
- Activity II. Clearly Defined, Answerable Study Question
- Activity III. Clearly Defined Study Indicator(s)
- Activity IV. Use a Representative and Generalizable Study Population
- Activity V. Valid Sampling Techniques (If Sampling was Used)
- Activity VI. Accurate/Complete Data Collection
- Activity VII. Appropriate Improvement Strategies
- Activity VIII. Sufficient Data Analysis and Interpretation
- Activity IX. Real Improvement Achieved
- Activity X. Sustained Improvement Achieved

All PIPs are scored as follows:

Met	(1) All critical elements were *Met*, **and** (2) 80 percent to 100 percent of all critical and non-critical elements were *Met*.
Partially Met	(1) All critical elements were *Met*, and 60 percent to 79 percent of all critical and non-critical elements were *Met*, **or** (2) One critical element or more was *Partially Met*.
Not Met	(1) All critical elements were *Met*, and <60 percent of all critical and non-critical elements were *Met*, **or** (2) One critical element or more was *Not Met*.
Not Applicable (N/A)	*N/A* elements (including critical elements if they were not assessed) were removed from all scoring.

For FY 06–07, the BHOs were provided an opportunity to resubmit additional information and/or documentation. The plans were required to take action for any evaluation element receiving a score of *Partially Met* or *Not Met*. The action could include resubmission of additional PIP documentation prior to final scoring. Future annual PIP submissions should include all information pertinent to the PIP study to achieve a *Met* status.

PIP Scores

For this PIP, HSAG reviewed Activities I through X. Table 2-1 and Table 2-2 show **ABC**'s scores based on HSAG's PIP evaluation of ***Improving Outcomes for High-Risk Youth Through AFFIRM Care Management***. Each activity has been reviewed and scored according to HSAG's validation methodology.

Table 2-1—FY 06-07 Performance Improvement Project Scores *for* Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

Review Activity	Total Possible Evaluation Elements (Including Critical Elements)	Total Met	Total Partially Met	Total Not Met	Total N/A	Total Possible Critical Elements	Total Critical Elements Met	Total Critical Elements Partially Met	Total Critical Elements Not Met	Total Critical Elements N/A
I. Appropriate Study Topic	6	6	0	0	0	1	1	0	0	0
II. Clearly Defined, Answerable Study Question	2	2	0	0	0	1	1	0	0	0
III. Clearly Defined Study Indicator(s)	7	6	0	0	1	3	3	0	0	0
IV. Use a Representative and Generalizable Study Population	3	3	0	0	0	2	2	0	0	0
V. Valid Sampling Techniques	6	0	0	0	6	1	0	0	0	1
VI. Accurate/Complete Data Collection	11	11	0	0	0	1	1	0	0	0
VII. Appropriate Improvement Strategies	4	3	0	0	1	No Critical Elements				
VIII. Sufficient Data Analysis and Interpretation	9	8	0	0	1	2	1	0	0	1
IX. Real Improvement Achieved	4	3	1	0	0	No Critical Elements				
X. Sustained Improvement Achieved	1	0	1	0	0	No Critical Elements				
Totals for All Activities	**53**	**42**	**2**	**0**	**9**	**11**	**9**	**0**	**0**	**2**

Table 2-2—FY 06-07 Performance Improvement Project Overall Score *for* Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

Percentage Score of Evaluation Elements Met*	**95%**
Percentage Score of Critical Elements Met**	**100%**
Validation Status***	***Met***

* The percentage score is calculated by dividing the total *Met* by the sum of the total *Met*, *Partially Met*, and *Not Met*.
** The percentage score of critical elements *Met* is calculated by dividing the total critical elements *Met* by the sum of the critical elements *Met*, *Partially Met*, and *Not Met*.
*** *Met* equals confidence/high confidence that the PIP was valid.
Partially Met equals low confidence that the PIP was valid.
Not Met equals reported PIP results that were not valid.

Validations and Findings Summary

This section summarizes the evaluation of the activities validated for the PIP. A description of the findings, strengths, requirements, and recommendations is outlined under each activity section. See Appendix B for a complete description of CMS rationale for each activity.

ABC's PIP evaluated quality of care and services. **ABC** used two study indicators to collect the data and assess the outcomes for this study. The study indicators measured inpatient readmissions at 30 days and reduction of CANS-MH domain scores. **ABC** completed 10 activities for this validation cycle.

Activity I. Appropriate Study Topic

Study Topic

ABC continued its study topic of ***Improving Outcomes for High-Risk Youth Through AFFIRM Care Management*** for the FY 06–07 validation cycle.

Finding(s)

Six of six evaluation elements, including one critical element, were *Met*.

Strength(s)

The study topic reflected a high-risk condition for **ABC** consumers. The goal was to reduce hospital readmissions and improve clinical and functional outcomes for high-risk consumers. The reduction in readmissions and improvements in outcomes for consumers would be the result of implementing a care management program called AFFIRM.

Requirement(s) (for Critical Elements)

There were no requirements identified for this activity during this review.

Recommendation(s) (for Noncritical Elements)

There were no recommendations identified for this activity during this review.

Activity II. Clearly Defined, Answerable Study Question

Study Question(s)

ABC's study question, as stated in its PIP Summary Form, was:

"Will consumer and family member participation in the AFFIRM Care Management Program for at least three months reduce inpatient psychiatric readmissions within 30 days for youths ages 0-17 and yield improvements in clinical and functional outcomes as measured by the CANS-MH assessment tool?"

Finding(s)

Both evaluation elements for this activity were *Met*, including the one critical element.

Strength(s)

The study question stated the problem in simple terms and set the focus of the study, which was to reduce psychiatric readmissions and improve clinical and functional outcomes for consumers. The goal of the study was to impact the quality of care provided to **ABC** consumers.

Requirement(s) (for Critical Elements)

There were no requirements identified for this activity during this review.

Recommendation(s) (for Noncritical Elements)

There were no recommendations identified for this activity during this review.

Activity III. Clearly Defined Study Indicator(s)

Study Indicator(s)

As stated in its PIP Summary Form, **ABC** had two study indicators:

- "Psychiatric inpatient readmission rates at 30 days."
- "Reduction of three- or six-month CANS-MH domain scores from initial CANS-MH domain scores."

Finding(s)

Six of seven evaluation elements were *Met*, including the three critical elements. One evaluation element was *Not Applicable* because the study indicators were not based on nationally recognized measures such as the Health Plan Employer Data and Information Set (HEDIS®)[1].

Strength(s)

The study indicators were developed to answer the study question and measure quality of care and services. The study indicators were well-designed to address CMS' requirements for evaluating quality.

Requirement(s) (for Critical Elements)

There were no requirements identified for this activity during this review.

Recommendation(s) (for Noncritical Elements)

There were no recommendations identified for this activity during this review.

Activity IV. Use a Representative and Generalizable Study Population

Study Population

The study population was different for each indicator. For the first indicator, the population consisted of consumers 0–17 years of age who had a psychiatric inpatient admission that met **ABC**'s defined criteria. The second indicator included a subset of the first indicator's population. Consumers in this subset were continuously enrolled and actively participating in the AFFIRM program for at least three months and had at least two CANS-MH assessments completed.

Finding(s)

All three evaluation elements, including the two critical elements, were *Met* for this activity.

Strength(s)

The study population was completely and thoroughly defined. The definition included requirements for the length of a consumer's enrollment and captured all consumers to whom the study question applied.

Requirement(s) (for Critical Elements)

There were no requirements identified for this activity during this review.

[1] **HEDIS®** refers to the Health Plan Employer Data and Information Set and is a registered trademark of the National Committee for Quality Assurance (NCQA).

Recommendation(s) (for Noncritical Elements)

There were no recommendations identified for this activity during this review.

Activity V. Valid Sampling Techniques

Sampling Technique(s)

The entire eligible population for each indicator was used. No sampling was performed.

Finding(s)

All six evaluation elements, including the one critical element, were *Not Applicable* for this activity.

Strength(s)

Because the entire eligible population for each indicator was used, no sampling was performed. The results of this study will represent all **ABC** consumers who met the eligible population criteria.

Requirement(s) (for Critical Elements)

There were no requirements identified for this activity during this review.

Recommendation(s) (for Noncritical Elements)

There were no recommendations identified for this activity during this review.

Activity VI. Accurate/Complete Data Collection

Data Collection

The data collection process included manual and administrative data sources. Manual data were collected from the CANS-MH Tool and administrative data from claim, encounter, and eligibility data. The process was appropriate for this study.

Finding(s)

Eleven of 11 evaluation elements, including the critical element related to manual data collection, were *Met* for this activity.

Strength(s)

The data collection techniques and processes used for this study were appropriate and well-implemented. The data elements were defined accurately and completely. The data sources were

identified and the processes used to collect the data were clear and easily understood. The manual data collection process used a tool that supported interrater reliability and ensured consistent and accurate data collection. The degree of administrative data completeness was estimated to be 98.99 percent.

Requirement(s) (for Critical Elements)

There were no requirements identified for this activity during this review.

Recommendation(s) (for Noncritical Elements)

There were no recommendations identified for this activity during this review.

Activity VII. Appropriate Improvement Strategies

Improvement Strategies

When implementing the AFFIRM program, **ABC** evaluated program effectiveness in reducing hospital readmissions. The program was aimed at providing enhanced clinical and social supports to children and adolescents whose recovery was complicated by multiple service systems and providers, lack of social supports, and lack of community connections.

Finding(s)

Three of four evaluation elements were *Met* for this activity. The other evaluation element related to standardizing and monitoring interventions was scored *Not Applicable*. **ABC** management determined that as a result of the deficiencies in the design of the AFFIRM intervention and the outcomes being measured, the program would no longer continue as originally designed.

Strength(s)

ABC performed causal and barrier analyses to identify possible interventions for the study. The interventions were related to data analyzed as part of the quality improvement process. The implemented interventions were likely to induce permanent change over time.

Requirement(s) (for Critical Elements)

There were no critical elements for this activity.

Recommendation(s) (for Noncritical Elements)

There were no recommendations identified for this activity during this review.

Activity VIII. Sufficient Data Analysis and Interpretation

Data Analysis and Interpretation

ABC provided data analysis and interpretation for baseline and two remeasurements for Study Indicator 1 and baseline and one remeasurement for Study Indicator 2. CANS-MH scores were collected upon enrollment and quarterly thereafter, and utilization data were collected annually. Data were analyzed annually.

Finding(s)

Eight of the nine evaluation elements were *Met* for this activity, including one critical element. One evaluation element was scored *Not Applicable* because sampling was not used for this study. This evaluation element was also a critical element.

Strength(s)

The data analysis was conducted according to the plan in the study. **ABC** completed statistical testing and provided *p* values for the differences between measurement periods. **ABC** presented the results in a clear and easily understood format and included a detailed interpretation of the data for each measurement period.

Requirement(s) (for Critical Elements)

There were no requirements identified for this activity during this review.

Recommendation(s) (for Noncritical Elements)

There were no recommendations identified for this activity during this review.

Activity IX. Real Improvement Achieved

Real Improvement Achieved

ABC provided statistical evidence that demonstrated statistically significant improvement in the rate for Study Indicator 1. Improvement in the rate for Study Indicator 2 was not statistically significant.

Finding(s)

Three of the four evaluation elements for this activity were *Met.* One evaluation element was *Partially Met* because the improvement in the rate for Study Indicator 2 was not statistically significant. There were no critical evaluation elements in this activity.

Strength(s)

Remeasurement methodology was the same as baseline methodology. There was documented improvement in outcomes of care, and the improvement appeared to be the result of the interventions. There was a large decline in the CANS-MH scores; however, **ABC** could not make any conclusive judgments about the success of the intervention due to the small number of consumers enrolled in the AFFIRM program in this measurement period.

Requirement(s) (for Critical Elements)

There were no critical elements for this activity.

Recommendation(s) (for Noncritical Elements)

There should be statistical evidence that observed improvement was true improvement for all of the study indicators.

Activity X. Sustained Improvement Achieved

Sustained Improvement Achieved

Although there was statistically significant improvement in readmission rates from baseline to the second remeasurement, **ABC** determined that the difference was clinically negligible due to the small number of consumers enrolled in the AFFIRM program.

Finding(s)

The one evaluation element for this activity received a *Partially Met* score.

Strength(s)

ABC will develop a new PIP focused on a high-priority area that needs improvement in the quality of mental health care, and that is better aligned to meet the needs of its consumers.

Requirement(s) (for Critical Elements)

There were no critical elements for this activity.

Recommendation(s) (for Noncritical Elements)

HSAG recommends for its new PIP that **ABC** target improvement in relevant areas of clinical care. The topic should be identified through data collection and analysis of comprehensive consumer needs, care, and services. The study topic should reflect the BHO's Medicaid enrollment in terms of demographic characteristics, prevalence of disease, and the potential consequences (risks) of disease. The study topic should reflect high-volume or high-risk conditions, address a broad

spectrum of care and services, and have the potential to affect consumer health, functional status, or satisfaction.

Section 4: Colorado FY 06-07 PIP Validation Tool: Improving Outcomes for High Risk Youth Through AFFIRM Care Management for Access Behavioral Care

DEMOGRAPHIC INFORMATION			
Health Plan Name:	Access Behavioral Care		
Study Leader Name:	Robert Bremer, MA, PhD	Title:	Behavioral Health Quality Program Manager
Phone Number:	(720) 744-5240	E-mail Address:	robert.bremer@coaccess.com
Name of Project/Study:	Improving Outcomes for High Risk Youth Through AFFIRM Care Management		
Type of Study:	Clinical		
Date of Study:	4/1/2006 to 2/1/2007		
Type of Delivery System:	BHO	Number of Medicaid Consumers in BHO: Number of Medicaid Consumers in Study:	8,121 2,999
Year 2 Validation	Initial Submission		

EVALUATION ELEMENTS		SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation			
I.	**Appropriate Study Topic: Topics selected for the study should reflect the Medicaid enrollment in terms of demographic characteristics, prevalence of disease, and the potential consequences (risks) of the disease. Topics could also address the need for a specific service. The goal of the project should be to improve processes and outcomes of health care. The topic may be specified by the State Medicaid agency or on the basis of Medicaid consumer input.**		
	1. Reflects high-volume or high-risk conditions (or was selected by the State). N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study topic reflected high-volume and high-risk conditions.
	2. Is selected following collection and analysis of data (or was selected by the State). N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study topic was selected following the collection and analysis of data.
	3. Addresses a broad spectrum of care and services (or was selected by the State). The scoring for this element will be Met or Not Met.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study topic addressed a broad spectrum of care and services.
	4. Includes all eligible populations that meet the study criteria. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study topic included all eligible populations that met the study criteria.
	5. Does not exclude consumers with special health care needs. The scoring for this element will be Met or Not Met.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	Consumers with special health care needs were not excluded.
C*	6. Has the potential to affect consumer health, functional status, or satisfaction. The scoring for this element will be Met or Not Met.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study topic had the potential to affect consumer health and functional status.

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS	SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation		

Results for Activity I				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
1	6	0	0	0

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS		SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation			
II.	**Clearly Defined, Answerable Study Question: Stating the study question(s) helps maintain the focus of the PIP and sets the framework for data collection, analysis, and interpretation.**		
	1. States the problem to be studied in simple terms. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study question stated the problem to be studied in simple terms.
C*	2. Is answerable. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study question was answerable.

Results for Activity II				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
1	2	0	0	0

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

	EVALUATION ELEMENTS	SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation			
III.	**Clearly Defined Study Indicator(s): A study indicator is a quantitative or qualitative characteristic or variable that reflects a discrete event (e.g., an older adult has not received a flu shot in the last 12 months) or a status (e.g., a consumer's blood pressure is or is not below a specified level) that is to be measured. The selected indicators should track performance or improvement over time. The indicators should be objective, clearly and unambiguously defined, and based on current clinical knowledge or health services research.**		
C*	1. Are well-defined, objective, and measurable. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study indicators were well-defined, objective, and measurable.
	2. Are based on current, evidence-based practice guidelines, pertinent peer review literature, or consensus expert panels.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study indicators were based on practice guidelines.
C*	3. Allow for the study question to be answered. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study indicators allowed for the study question to be answered.
	4. Measure changes (outcomes) in health or functional status, consumer satisfaction, or valid process alternatives. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study indicators measured changes in consumer health and functional status.
C*	5. Have available data that can be collected on each indicator. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	There were available data collected on each study indicator.
	6. Are nationally recognized measures such as HEDIS specifications, when appropriate. The scoring for this element will be Met or N/A.	☐ Met ☐ Partially Met ☐ Not Met ☑ N/A	The study indicators were not nationally recognized measures.
	7. Includes the basis on which the indicator(s) was adopted, if internally developed.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The basis on which each indicator was adopted was provided.

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS	SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation		

Results for Activity III				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
3	6	0	0	1

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS		SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation			
IV.	**Use a representative and generalizable study population: The selected topic should represent the entire eligible Medicaid enrollment population with systemwide measurement and improvement efforts to which the PIP study indicators apply.**		
C*	1. Is accurately and completely defined. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study population was accurately and completely defined.
	2. Includes requirements for the length of a consumer's enrollment in the BHO.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	Requirements for length of a consumer's enrollment were included.
C*	3. Captures all consumers to whom the study question applies. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The study population captured all consumers to whom the study question applied.

Results for Activity IV				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
2	3	0	0	0

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

	EVALUATION ELEMENTS	SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation			
V.	**Valid Sampling Techniques: (This activity is only scored if sampling was used.) If sampling is to be used to select consumers of the study, proper sampling techniques are necessary to provide valid and reliable information on the quality of care provided. The true prevalence or incidence rate for the event in the population may not be known the first time a topic is studied.**		
	1. Consider and specify the true or estimated frequency of occurrence.	☐ Met ☐ Partially Met ☐ Not Met ☑ N/A	Sampling was not used.
	2. Identify the sample size.	☐ Met ☐ Partially Met ☐ Not Met ☑ N/A	Sampling was not used.
	3. Specify the confidence level.	☐ Met ☐ Partially Met ☐ Not Met ☑ N/A	Sampling was not used.
	4. Specify the acceptable margin of error.	☐ Met ☐ Partially Met ☐ Not Met ☑ N/A	Sampling was not used.
C*	5. Ensure a representative sample of the eligible population.	☐ Met ☐ Partially Met ☐ Not Met ☑ N/A	Sampling was not used.
	6. Are in accordance with generally accepted principles of research design and statistical analysis.	☐ Met ☐ Partially Met ☐ Not Met ☑ N/A	Sampling was not used.

Results for Activity V				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
1	0	0	0	6

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

	EVALUATION ELEMENTS	SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation			
VI.	**Accurate/Complete Data Collection: Data collection must ensure that the data collected on the PIP indicators are valid and reliable. Validity is an indication of the accuracy of the information obtained. Reliability is an indication of the repeatability or reproducibility of a measurement.**		
	1. Clearly defined data elements to be collected. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The data elements collected were identified.
	2. Clearly identified sources of data. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The sources of data were identified.
	3. A clearly defined and systematic process for collecting data that includes how baseline and remeasurement data will be collected. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The process for collecting data was defined and systematic.
	4. A timeline for the collection of baseline and remeasurement data. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	A timeline for the collection of data was included.
	5. Qualified staff and personnel to abstract manual data.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	Qualified staff members were used to abstract manual data.
C*	6. A manual data collection tool that ensures consistent and accurate collection of data according to indicator specifications.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The manual data collection tool ensured consistent and accurate collection of data.
	7. A manual data collection tool that supports interrater reliability.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The manual data collection tool supported interrater reliability.
	8. Clear and concise written instructions for completing the manual data collection tool.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The instructions for completing the manual data collection tool were clear and concise.
	9. An overview of the study in written instructions.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	An overview of the study was included in the written instructions for the manual data collection tool.

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS		SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation			
VI.	**Accurate/Complete Data Collection: Data collection must ensure that the data collected on the PIP indicators are valid and reliable. Validity is an indication of the accuracy of the information obtained. Reliability is an indication of the repeatability or reproducibility of a measurement.**		
	10. Administrative data collection algorithms/flow charts that show activities in the production of indicators.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	A description of the administrative data collection process was included.
	11. An estimated degree of administrative data completeness. Met = 80 - 100% Partially Met = 50 - 79% Not Met = <50% or not provided	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The estimated degree of administrative data completeness was reported as 98.99 percent.

Results for Activity VI				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
1	11	0	0	0

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS		SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation			
VII.	**Appropriate Improvement Strategies: Real, sustained improvements in care result from a continuous cycle of measuring and analyzing performance, and developing and implementing systemwide improvements in care. Interventions are designed to change behavior at an institutional, practitioner, or consumer level.**		
	1. Related to causes/barriers identified through data analysis and quality improvement processes. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The interventions were related to causes/barriers identified through data analysis and quality improvement processes.
	2. System changes that are likely to induce permanent change.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The interventions were system changes that were likely to induce permanent change.
	3. Revised if the original interventions were not successful.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	ABC management determined it was no longer possible to continue the AFFIRM program as originally designed. Resources will be refocused to higher priority areas that serve a greater number of consumers.
	4. Standardized and monitored if interventions were successful.	☐ Met ☐ Partially Met ☐ Not Met ☑ N/A	There was improvement in outcomes of CANS-MH scores and improvement in the rate of readmissions; however, there was a small number of consumers enrolled in the AFFIRM program for the current measurement period. ABC management determined that the AFFIRM program will no longer continue as originally designed.

Results for Activity VII				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
0	3	0	0	1

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS			SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation				
VIII.	**Sufficient Data Analysis and Interpretation: Describe the data analysis process on the selected clinical or nonclinical study indicators. Include the statistical analysis techniques used.**			
C*	1.	Is conducted according to the data analysis plan in the study design. N/A is not applicable to this element for scoring.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The data analysis was conducted according to the data analysis plan.
C*	2.	Allows for the generalization of results to the study population if a sample was selected. If no sampling was performed, this element is scored N/A.	☐ Met ☐ Partially Met ☐ Not Met ☑ N/A	A sample was not selected.
	3.	Identifies factors that threaten internal or external validity of findings.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	Factors that threatened the internal or external validity of findings were identified.
	4.	Includes an interpretation of findings.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	An interpretation of findings was included.
	5.	Is presented in a way that provides accurate, clear, and easily understood information.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The information was presented in an accurate and easily understood way.
	6.	Identifies initial measurement and remeasurement of study indicators.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	Baseline and two remeasurements were provided for Study Indicator 1. Data for Study Indicator 2 was only collected for Baseline and Remeasurement 1.
	7.	Identifies statistical differences between initial measurement and remeasurement.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	Statistical differences between measurements were identified.
	8.	Identifies factors that affect the ability to compare initial measurement with remeasurement.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	Factors that affected the ability to compare measurements were identified.
	9.	Includes interpretation of the extent to which the study was successful.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	An interpretation of the extent to which the study was successful was included.

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS	SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation		

Results for Activity VIII				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
2	8	0	0	1

* "C" in this column denotes a critical evaluation element.

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS		SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation			
IX.	**Real Improvement Achieved: Describe any meaningful change in performance observed and demonstrated during baseline measurement. Discuss any random year-to-year variation, population changes, and sampling error that may have occurred during the measurement process.**		
	1. Remeasurement methodology is the same as baseline methodology.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	Remeasurement methodology was the same as baseline methodology.
	2. There is documented improvement in processes or outcomes of care.	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	There was documented improvement in outcomes of care.
	3. The improvement appears to be the result of planned intervention(s).	☑ Met ☐ Partially Met ☐ Not Met ☐ N/A	The improvement appeared to be the result of the planned interventions.
	4. There is statistical evidence that observed improvement is true improvement.	☐ Met ☑ Partially Met ☐ Not Met ☐ N/A	The improvement in the rate of readmissions was statistically significant. The improvement in CANS-MH scores was not statistically significant.

Results for Activity IX				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
0	3	1	0	0

** This number is a tally of the total number of critical evaluation elements for this review activity.

EVALUATION ELEMENTS	SCORING	COMMENTS
Performance Improvement Project/Health Care Study Evaluation		
X. **Sustained Improvement Achieved: Describe any demonstrated improvement through repeated measurements over comparable time periods. Discuss any random year-to-year variation, population changes, and sampling error that may have occurred during the remeasurement process.**		
1. Repeated measurements over comparable time periods demonstrate sustained improvement, or that a decline in improvement is not statistically significant.	☐ Met ☑ Partially Met ☐ Not Met ☐ N/A	Although there was statistically significant improvement in readmission rates from Baseline to Remeasurement 2, ABC determined the difference was clinically negligible. ABC completed three measurements of Study Indicator 1 and two measurements of Study Indicator 2. The baseline rate for Study Indicator 1 (rate of readmissions) was 5.6 percent and the benchmark was 12 percent. ABC determined that Study Indicator 1 was not the best indicator to use because of the small number of consumers enrolled in the AFFIRM program and the baseline rate was well below the benchmark. ABC will develop a new PIP where resources can be better targeted to achieve positive outcomes for consumers.

Results for Activity X				
# of Elements				
Critical Elements**	**Met**	**Partially Met**	**Not Met**	**Not Applicable**
0	0	1	0	0

** This number is a tally of the total number of critical evaluation elements for this review activity.

Table A-1—FY 06-07 PIP Validation Report Scores: Improving Outcomes for High Risk Youth Through AFFIRM Care Management for Access Behavioral Care											
Review Activity		**Total Possible Evaluation Elements (Including Critical Elements)**	**Total Met**	**Total Partially Met**	**Total Not Met**	**Total N/A**	**Total Possible Critical Elements**	**Total Critical Elements Met**	**Total Critical Elements Partially Met**	**Total Critical Elements Not Met**	**Total Critical Elements N/A**
I.	Appropriate Study Topic	6	6	0	0	0	1	1	0	0	0
II.	Clearly Defined, Answerable Study Question	2	2	0	0	0	1	1	0	0	0
III.	Clearly Defined Study Indicator(s)	7	6	0	0	1	3	3	0	0	0
IV.	Use a representative and generalizable study population	3	3	0	0	0	2	2	0	0	0
V.	Valid Sampling Techniques	6	0	0	0	6	1	0	0	0	1
VI.	Accurate/Complete Data Collection	11	11	0	0	0	1	1	0	0	0
VII.	Appropriate Improvement Strategies	4	3	0	0	1	0	No Critical Elements			
VIII.	Sufficient Data Analysis and Interpretation	9	8	0	0	1	2	1	0	0	1
IX.	Real Improvement Achieved	4	3	1	0	0	0	No Critical Elements			
X.	Sustained Improvement Achieved	1	0	1	0	0	0	No Critical Elements			
Totals for All Activities		**53**	**42**	**2**	**0**	**9**	**11**	**9**	**0**	**0**	**2**

Table A-2—FY 06-07 PIP Validation Report Overall Scores: Improving Outcomes for High Risk Youth Through AFFIRM Care Management for Access Behavioral Care	
Percentage Score of Evaluation Elements Met*	**95%**
Percentage Score of Critical Elements Met**	**100%**
Validation Status***	***Met***

* The percentage score is calculated by dividing the total Met by the sum of the total Met, Partially Met, and Not Met.

** The percentage score of critical elements Met is calculated by dividing the total critical elements Met by the sum of the critical elements Met, Partially Met, and Not Met.

*** Met equals confidence/high confidence that the PIP was valid.
Partially Met equals low confidence that the PIP was valid.
Not Met equals reported PIP results that were not credible.

EVALUATION OF THE OVERALL VALIDITY AND RELIABILITY OF PIP/STUDY RESULTS
HSAG assessed the implications of the study's findings on the likely validity and reliability of the results based on CMS protocols. HSAG also assessed whether the State should have confidence in the reported PIP findings. Determining when an accumulation of threats to validity and reliability, and PIP design problems, reach a point at which the PIP findings are no longer credible is always a judgment call.

***Met** = Confidence/high confidence in reported PIP results

****Partially Met** = Low confidence in reported PIP results

*****Not Met** = Reported PIP results not credible

Summary of Aggregate Validation Findings

*		Met	**		Partially Met	***		Not Met
	X							

Summary statement on the validation findings:

Activities I through X were assessed for this PIP Validation Report. Based on the validation of this PIP study, HSAG's assessment determined high confidence in the results.

HSAG HEALTH SERVICES ADVISORY GROUP

Appendices

for Access Behavioral Care

Introduction

The appendices consist of documentation supporting the validation process conducted by HSAG using the CMS Protocol for validating PIPs. Appendix A is the study submitted to HSAG for review, Appendix B is CMS rationale for each activity, and Appendix C includes PIP definitions and explanations.

- Appendix A: **Access Behavioral Care**'s PIP Study: ***Improving Outcomes for High-Risk Youth Through AFFIRM Care Management***
- Appendix B: CMS Rationale by Activity
- Appendix C: Definitions and Explanations by Activity

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

DEMOGRAPHIC INFORMATION			
BHO Name and ID:	Access Behavioral Care		
Study Leader Name:	Robert Bremer, MA, PhD	Title:	Behavioral Health Quality Program Manager
Telephone Number:	720-744-5240	E-mail Address:	robert.bremer@coaccess.com
Name of Project/Study:	Improving Outcomes For High-Risk Youth Through AFFIRM Care Management		
Type of Study:	☒ Clinical ☐ Nonclinical		
Date of Study Period:	~~April 1, 2005 to February 1, 2006~~ April 1, 2006, February 1, 2007		

Number of Medicaid Consumers served by BHO: 8,121
(Total all ages FY06 based on paid claims)

Number of Medicaid Consumers served by BHO: 2,999
(Total youth ages 0-17 FY06 based on paid claims)

Number of Medicaid Consumers in Project/Study: Calendar Year 2006: 5
Calendar Year 2005: 33

Section to be completed by HSAG

_____ Year 1 Validation _____ Initial Submission _____ Resubmission

__X__ Year 2 Validation __X__ Initial Submission _____ Resubmission

_____ Year 3 Validation _____ Initial Submission _____ Resubmission

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

A. Activity I: Choose the Selected Study Topic. Topics selected for study should reflect the Medicaid enrollment in terms of demographic characteristics, prevalence of disease, and the potential consequences (risks) of the disease. Topics could also address the need for a specific nonclinical service. The goal of the project should be to improve processes and outcomes of health care for the full affected population. The topic may be specified by the State Medicaid agency or on the basis of Medicaid consumer input.

Study Topic: The purpose of the study is to determine whether participation in a structured care management program can reduce psychiatric hospital readmissions and improve clinical and functional outcomes for high-risk child and adolescent consumers.

Colorado Access began developing care management and disease management programs in response to analyses of diagnostic data indicating that adult members with a chronic medical condition and co-morbid depression, anxiety, or substance abuse had higher utilization of general medical services. The presence of a mental health diagnosis was found to increase total health care costs by a factor of 2.24[1]. "Usual care" for this population of complex Medicaid members typically consisted of either treatment solely in the primary care setting, or by multiple providers who did not adequately communicate or coordinate with one another, often leading to poorer clinical outcomes. To address quality of care and costs, with both physical and behavioral health plan resources, Colorado Access has expanded care management beyond a traditional medical diagnosis-specific focus to include behavioral health care components and programs.

A resource paper recently published by the Center for Health Care Strategies notes that states and public sector purchasers are beginning to recognize the potential benefits of care management programs for behavioral health disorders and implementing programs or pilots[2]. Although evaluation of care management programs to date has been primarily limited to those programs that have been employed in the primary care setting, research suggests that behavioral health care management programs can be effective in reducing or eliminating duplicated services[3], improving medication compliance and management of side effects, minimizing high-risk behaviors, decreasing the use of intensive services, increasing consumer and provider satisfaction, and achieving net savings[4]. Colorado Access has found encouraging trends in its implementation of a care management program to improve the primary care treatment of depression in adult members with co-existing medical conditions. Preliminary analysis of data indicates clinical improvement in depression severity, decreases in emergency room and inpatient utilization, and an overall savings in medical costs[5].

While the care management programs that Colorado Access has created have common structural elements (i.e., risk stratification methodology, standard assessments, care management protocols, collaborative care plans, consumer and provider tool kits, and analytic tools to track outcomes and return on investment), further analysis of member and claims data specific to Access Behavioral Care (ABC) suggested a different approach to child and adolescent care management, for which little if any research literature exists. Extensive review of data revealed that diagnosis was not a reliable

[1] Thomas, M., Waxmonsky, J., McGinnis, G., Gabow, P., Socherman, R., Rost, K. (2005). Prevalence of psychiatric disorders and costs of care among adult enrollees in a Medicaid HMO, 56(11): 1394-401.
[2] Gelber, S., Dougherty, R. (2005). Disease management for chronic behavioral health and substance use disorders. Center for Health Care Strategies resource paper.
[3] Badger, L., et. al. (1999). Management of mental disorders in rural primary care: a proposal for integrated psychosocial services. Journal of Family Practice, 48(10): 813-818.
[4] Katon, W. et al. (1996). A multifaceted intervention to improve treatment of depression in primary care. Archive of General Psychiatry, 53(10), 924-932.
[5] Thomas, M., Waxmonsky, J., McGinnis, G., Barry, C. (2006). Realigning clinical and economic incentives to support depression management within a Medicaid population: the Colorado Access experience. Administration and Policy in Mental Health, 33(1).

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

A. Activity I: Choose the Selected Study Topic. Topics selected for study should reflect the Medicaid enrollment in terms of demographic characteristics, prevalence of disease, and the potential consequences (risks) of the disease. Topics could also address the need for a specific nonclinical service. The goal of the project should be to improve processes and outcomes of health care for the full affected population. The topic may be specified by the State Medicaid agency or on the basis of Medicaid consumer input.

indicator for identifying care management candidates in this population. For example, an analysis of 156 adolescent consumers ages 12-17 diagnosed with bipolar disorder indicated that only 34 had net medical costs over $10,000, and of those, 15 would have been ineligible for a bipolar disorder care management program due to aging out or RTC placement.

A subsequent broader analysis of the top 250 highest-cost ABC consumers under 18 years of age, regardless of diagnosis, resulted in a finding that Access Behavioral Care children and adolescents identified as high-cost and high-risk, and consequently, as potential candidates for care management, had a high likelihood of multiple system involvement, meaning at least one of the following in addition to mental health treatment:

- Substance abuse treatment
- Child Welfare system involvement
- Special education
- Medical co-morbidity
- Developmental disabilities
- Juvenile justice system involvement

These high-risk consumers with multiple and special health care needs may be more likely to require psychiatric hospitalization, out-of-home placement, or treatment in more restrictive settings. Fiscal year 2004 data indicates that consumers ages 0-17, many of whom meet the AFFIRM criterion of multi-system involvement, comprised over 45% of all psychiatric hospital admissions. Of these, 14% were readmitted within 30 days.

Based on the hypothesis that these high-risk children and adolescents could be proactively identified through provider requests for authorization of higher-level services, Access Behavioral Care designed a consumer- and family-focused approach to risk assessment, early intervention, and the coordination of care with any involved organizations. This care management program for high-risk children and adolescents has been named AFFIRM: Access Family-Focused Intervention and Recovery Model. All children and adolescents who require a level of care more intensive than routine outpatient are screened. If they meet criteria, they and their family members are engaged through outreach and enrolled. Specific areas of intervention are defined through structured assessment and as warranted include wellness and illness self-management education, linkage to community resources and supports, referrals to other needed services, advocacy and assistance by resource coordinators, and coordination of care across multiple delivery systems. Information and psychiatric consultation is made available to providers of care. Individualized intervention strategies are guided by BHO-level care plans developed in collaboration with consumers, their family members, and multi-system providers.

Incentives for enrolled consumers and families include increased stability, reduction in mental health symptoms, and improved functioning through coordinated attention to their psychiatric, medical, and psychosocial needs. The aims of AFFIRM care management are to shift patterns of utilization to the most clinically appropriate least restrictive settings, improve clinical and functional outcomes, improve consumer and family quality of life and satisfaction, and reduce costs of care. For purposes of this PIP, evaluation will focus on hospital readmissions and clinical/ functional outcomes.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

B. Activity II: The Study Question. Stating the question(s) sets the framework for data collection, analysis, and interpretation.
Study Question: Will consumer and family member participation in the AFFIRM Care Management Program for at least three months reduce inpatient psychiatric readmissions within 30 days for youth ages 0-17 and yield improvements in clinical and functional outcomes as measured by the Child and Adolescent Needs and Strengths – Mental Health* (CANS-MH) assessment tool? * Developed by John Lyons, Ph.D., Director, Mental Health Services and Policy Program, Northwestern University, Chicago, IL

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

C. Activity III: Selected Study Indicators. A study indicator is a quantitative or qualitative characteristic or variable that reflects a discrete event (e.g., rates of hospital readmissions within 30 or 90 days), or a status (e.g., percent of consumers reporting that they actively participate in treatment planning) that is to be measured. The selected indicators should be appropriate for the study topic and question as well as track performance or improvement over time. The indicators should be objective, clearly and unambiguously defined, and based on current clinical knowledge or health services research.

Study Indicator #1:	Psychiatric inpatient readmission rates at 30 days
Numerator:	Number of inpatient admissions for treatment of a primary covered mental health diagnosis (psychiatric hospital or acute treatment unit; emergency department visits excluded) within 30 days of a date of discharge (excluding transfer to a hospital-based facility), following date of enrollment, for consumers ages 0-17 enrolled in AFFIRM ≥ 3 months
Denominator:	Number of psychiatric inpatient admissions for treatment of a primary covered mental health diagnosis, following date of enrollment, for all consumers ages 0-17 enrolled in AFFIRM ≥ 3 months
Measurement Period Dates:	April 1,through December 31[6] of the measurement year
Benchmark:	12.0%: mean 30-day readmission rate
Source of Benchmark:	Children's Mental Health Benchmarking Project, Center for Health Care Strategies, July 2003
Baseline Goal:	30-day readmission rate less than or equal to Children's Mental Health Benchmarking Project mean

[6] Although the measurement period extends until 12/31, members are only identified through 10/31 to allow to claims lag.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

C. Activity III: Selected Study Indicators. A study indicator is a quantitative or qualitative characteristic or variable that reflects a discrete event (e.g., rates of hospital readmissions within 30 or 90 days), or a status (e.g., percent of consumers reporting that they actively participate in treatment planning) that is to be measured. The selected indicators should be appropriate for the study topic and question as well as track performance or improvement over time. The indicators should be objective, clearly and unambiguously defined, and based on current clinical knowledge or health services research.

Study Indicator #2:	Reduction of three- or six-month CANS-MH domain scores from initial CANS-MH domain scores
Numerator:	Number of consumers ages 0-17 enrolled in the AFFIRM program with four of six CANS-MH domain scores at three or six months (as determined by length of enrollment) lower than initial CANS-MH domain scores
Denominator:	Number of consumers ages 0-17 enrolled in the AFFIRM program with CANS-MH domain scores at three or six months
Measurement Period Dates:	April 1 through February 1 of the measurement year
Benchmark:	No benchmark available
Source of Benchmark:	N/A
Baseline Goal:	60% of consumers ages 0-17 enrolled in the AFFIRM program show reduction of one point or more in four of six CANS-MH domain scores at three or six months (as determined by length of enrollment)

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

D. Activity IV: Identified Study Population. The study population should be clearly defined to represent the entire population to which the PIP study question and indicators apply. The length of consumer enrollment should be considered and defined. All selection criteria should be listed here. Once the population is identified, a decision must be made whether to review data for the entire population or a sample of that population.

Identified Study Population:

For indicator #1 regarding inpatient readmission rates, the study population consists of ABC youth ages 0-17 meeting the following criteria:

- A provider request for authorization of Level 3 (psychiatric inpatient or acute treatment unit; emergency department visits are excluded since no authorization is required) or Level 2 (residential, partial hospitalization, day treatment, or home-based) services has been received by ABC, beginning April 1st of the measurement year;
- Residential Treatment Center (RTC) placement paid for by Department of Human Services, or Division of Youth Corrections (DYC) commitment, is not current or imminent;
- Screening conducted with the Risk and Stability Index (RSI) upon admission to a Level 2 service or discharge from a Level 3 service yields a minimum score of two in at least three of the eight domains;
- The family has agreed to be enrolled in the AFFIRM program;
- The consumer has been continuously enrolled with no gaps in enrollment to ensure that ABC has remained the payer of services delivered to the consumer; and
- The consumer and family remained active in the program for a period of at least three months.

For indicator #2 regarding reduction in CANS-MH domain scores, the study population consists of a subset of AFFIRM program participants who met the above criteria, were enrolled and actively participating in the AFFIRM program for a period of at least three months, and had at least two CANS-MH assessments completed (initial assessment and at least one follow-up assessment).

Of those meeting the criteria, the entire population was included in the study and no sampling was conducted.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

E. Activity V: Sampling Methods. If sampling is to be used to select consumers of the study, proper sampling techniques are necessary to provide valid and reliable information on the quality of care provided. The true prevalence or incidence rate for the event in the population may not be known for the first time a topic is studied. In this case, an estimate should be used and the basis for that estimate indicated.

Measure	Sample Error and Confidence Level	Sample Size	Population	Method for Determining Size (*describe*)	Sampling Method (*describe*)
Not applicable – total eligible population used					

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

F. Activity VIa: Data Collection Procedures. Data collection must ensure that the data collected on the PIP indicators are valid and reliable. Validity is an indication of the accuracy of the information obtained. Reliability is an indication of the repeatability or reproducibility of a measurement.

Data Sources

[☐] Hybrid (medical/treatment records and administrative)

[☐] Medical/treatment record abstraction
Record Type
[☐] Outpatient
[☐] Inpatient
[☐] Other

[☐] Other data

Description of Data Collection Staff:

Service Coordinators: Licensed mental health professionals responsible for day-to-day utilization management and Care Management activities, with an average of nine years licensure and five years employment with Colorado Access. Service Coordinators are responsible for completing the Problem Presentation, Risk Behaviors, and Strengths sections of the CANS-MH.

Resource Coordinators: Paraprofessional staff who are either mental health consumers well along in their recovery, or family members of consumers. These staff have an intimate knowledge of the mental health and human services system from their own and their family member's experiences, and have an average of five years of employment with Colorado Access. Resource Coordinators are responsible for completing the Functioning, Care Intensity and

[☒] Administrative data
Data Source
[☒] Programmed pull from claims/encounters
[☐] Complaint/appeal
[☐] Pharmacy data
[☐] Telephone service data /call center data
[☐] Appointment/access data
[☐] Delegated entity/vendor data _____
[☒] Other: Eligibility data
[☒] Other: Colorado Access Care Management (CACM) database records
[☒] Other: Child and Adolescent Needs & Strengths – Mental Health (CANS-MH) assessment

Other Requirements
[☒] Data completeness assessment attached (Attachment 1
[☒] Coding verification process attached (Attachment 2)
[☒] Data collection tool attached (Attachment 3)
[☒] Data collection instructions attached (Attachment 4)
[☒] Summary of data collection training attached (Attachment 5
[☒] IRR process and results attached (Attachment 5)

[☐] Survey Data
Fielding Method
[☐] Personal interview
[☐] Mail
[☐] Phone with CATI script
[☐] Phone with IVR
[☐] Internet
[☐] Other _____

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

Organization, and Caregiver Needs and Strengths sections of the CANS-MH.	Other Requirements [☐] Number of waves _____ [☐] Response rate _____ [☐] Incentives used _____

F. Activity VIb: Data Collection Cycle.	Data Analysis Cycle.
[☒] Once a year [☐] Twice a year [☐] Once a season [☐] Once a quarter [☐] Once a month [☐] Once a week [☐] Once a day [☐] Continuous [☐] Other (list and describe): CANS-MH scores are collected upon enrollment and quarterly thereafter, and retrieved for data analysis annually. Utilization data is collected annually.	[☒] Once a year [☐] Once a season [☐] Once a quarter [☐] Once a month [☐] Continuous [☐] Other (list and describe):

F. Activity VIc. Data Analysis Plan and Other Pertinent Methodological Features

Data analysis is conducted in the February following the measurement period, prior to report submission.

For study indicator #1, analysis is conducted annually for consumers with at least three months active enrollment in the AFFIRM program, on claims received for dates of service April 1st through December 31st of the measurement year. To allow for a three-month lag for claim run-out and maximize data completeness members are only identified through October 31st. For the current submission, data completeness was estimated to be 98.99% (Attachment 1). Inpatient admissions claims data are extracted from PowerSTEPP, the Colorado Access transaction system, using administrative methodology. This information is downloaded daily into Business Objects, a data analysis software program and decision support tool used by Colorado Access.

Service codes and description, place of service, and admission and discharge dates for all inpatient psychiatric hospitalization or acute treatment unit admissions for treatment of a primary mental health diagnosis are extracted for all enrolled consumers. Both paid and denied claims are

F. Activity VIc. Data Analysis Plan and Other Pertinent Methodological Features

included. A count of all discrete admissions for each consumer following the date of his or her enrollment in the AFFIRM program constitutes the denominator. Multiple admissions following the date of the consumer's enrollment in the program are identified, and each readmission falling within

30 days of the discharge date of a prior admission is counted in the numerator. An admission that results in a direct transfer to another acute care facility is counted as one admission in either the numerator or denominator. Duplicate claims that were not paid are excluded.

Baseline for the initial report was established by measuring the 30-day inpatient readmission rate, using the above specifications, for the year prior to program start date, to compare pre-enrollment and enrollment phase utilization. A comparison of consumers who were screened and risked-in by virtue of RSI scores but whose families declined to participate in the AFFIRM program, withdrew participation after enrollment in the program, lost Access Behavioral Care eligibility, or were wait-listed due to limited resource availability. For both groups, psychiatric inpatient admissions claims data with dates of services from April 1st through October 31st of the measurement year will be extracted and analyzed using the methodology described in the preceding paragraph.

For study indicator #2, analysis is conducted annually on all initial and follow-up CANS-MH scores obtained through February 1st of the measurement year to extend opportunities for completion of repeat CANS-MH assessments for those enrolled in the AFFIRM program. The CANS-MH is scheduled to be completed upon consumer and family enrollment, and at quarterly intervals. ~~Enrolled consumers may not all have an initial or follow-up CANS-MH assessment, depending upon date of enrollment in the program and degree of difficulty in connecting with the family and/or provider.~~ **The collection of the CANS-MH data proved to be overly burdensome to staff and families. Difficulties with the data collection will be discussed in detail in Activitys 8B and 10.**

The CANS-MH is a standardized instrument developed in 1999 with demonstrated reliability and validity as a prospective and retrospective assessment tool. The manual for administration, which contains clearly defined coding criteria (Attachment 4, as well as an interviewing guide and glossary of terms (Attachment xx), were provided to Care Management staff. Training on the use of the CANS-MH was conducted over the course of three sessions, and included inter-rater reliability studies using three case vignettes (Attachment 5. Results were shared and discussed within the group as a learning opportunity to increase accuracy of ratings. Additional meetings were held to educate staff about the AFFIRM program in general, the design of the Performance Improvement Project and the role of the CANS-MH in the evaluation of the program.

F. Activity VIc. Data Analysis Plan and Other Pertinent Methodological Features

As noted in the section above under Data Sources, clinically qualified Service Coordinators complete the Problem Presentation, Risk Behaviors, and Strengths sections of the CANS-MH, and experienced Resource Coordinators complete the Functioning, Care Intensity and Organization, and Caregiver Needs and Strengths sections of the CANS-MH for their assigned cases. Each dimension is rated on a 4-point scale, based on information received from the family and/or provider, and/or review of case files, reflecting the consumer and family's current status. A score of "0" on any particular item indicates no evidence and no need for action, a score of "1" indicates present but mild and suggests "watchful waiting" to see whether action is needed or prevention planning, a score of "2" indicates present to a moderate degree and a need for action, and a score of "3" indicates present to a severe or profound degree and a need for either immediate or intensive action. Therefore, in all cases, a low rating is positive, and a change in score of one point reflects a qualitatively pronounced difference in consumer and/or family status. Ratings are entered into CACM by the Service Coordinator or Resource Coordinator who conducts the assessment.

For the analysis, CANS-MH scores are extracted from the CACM database for enrolled consumers having an initial CANS-MH assessment completed upon enrollment and at least one follow-up CANS-MH assessment at three or six months following enrollment. Item scores for each of the six domains (described in Attachment 4 are summed to yield a domain score. Comparisons are then made between the initial CANS-MH assessment domain scores, which constitute the baseline scores, and the most recent follow-up CANS-MH assessment domain scores, which constitute the remeasurement scores. The number of domains for each consumer that are one point or more lower upon remeasurement than the initial measurement is counted. The number of consumers who obtain four of six domain scores at least one point lower upon remeasurement than the initial measurement is counted in the numerator, the total number of consumers with at least one follow-up CANS-MH assessment is counted in the denominator, and the percentage of consumers obtaining lower scores in at least four of six domains upon remeasurement is calculated. Since the CANS-MH is administered only to enrolled consumers and families, no comparison group is available.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

G. Activity VII. Improvement Strategies. Real, sustained improvements in care result from a continuous cycle of measuring and analyzing performance, and developing and implementing system-wide improvements in care. Describe interventions designed to change behavior at an institutional, practitioner, or consumer level.

Describe interventions.

Baseline to Remeasurement 1

The Access Family Focused Intervention and Recovery Model (AFFIRM) is a care management program aimed at providing enhanced clinical and social supports to child and adolescent consumers whose path to recovery is complicated by involvement with multiple service systems, multiple providers, lack of social supports, and lack of community connections. These consumers may also have family members with significant mental health, physical health, developmental, legal, or substance abuse concerns, impacting their ability to benefit from services. The approach includes the proactive identification of youth with the highest levels of co-morbidity, acuity, and service utilization, and early intervention with the family and any involved systems. A primary focus is the mobilization and integration of treatment and community resources through care coordination activities among multiple providers and human service agencies, facilitating the consumers' stabilization into less intensive levels of care. Once stabilized, consumers continue to be followed to maintain stability and promote the maximum level of recovery.

The AFFIRM Care Management program, as described in the Operational Manual (Attachment 6) was implemented in April 2005. Following identification of high-risk consumers through service authorizations for Levels 2 and 3, administration of the Risk and Stability Index (RSI – Operational Manual, page 21) by ABC clinical staff, and determination of eligibility for the program, the process of outreach is initiated for those consumers who meet criteria for enrollment in the program. An assigned Outreach Specialist reviews the RSI and CareSTEPP records (Colorado Access's utilization management software), and conducts an outreach telephone call to the family within two business days of completion of the RSI. A standardized script is utilized (Operational Manual, pages 10-11). The goal is to engage the family, gather information about services the consumer is receiving and potential areas of need, and determine the consumer's and/or family's willingness to meet with care management staff and participate in the program. If telephone outreach cannot be completed, alternative outreach strategies are planned, such as enlisting the consumer's provider for consultation and assistance or meeting the consumer and family at the provider site during a scheduled appointment. All attempted and completed contacts are documented in the Colorado Access Care Management (CACM) database.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

G. Activity VII. Improvement Strategies. Real, sustained improvements in care result from a continuous cycle of measuring and analyzing performance, and developing and implementing system-wide improvements in care. Describe interventions designed to change behavior at an institutional, practitioner, or consumer level.

Families who decline participation are given the ABC contact number and encouraged to call if needs arise in the future, and the case is closed. Those who feel they could benefit from additional support or services and are interested in meeting with a Resource Coordinator are enrolled in the AFFIRM program. Each enrolled consumer has an identified Resource Coordinator that maintains their relationship with the consumer and family longitudinally to ensure continuity of care and a clearly identified and accessible advocate for their needs. Service Coordinators are assigned according to the consumer's provider. The Resource Coordinator and Service Coordinator dyad then begins to collect data for completion of the CANS-MH assessment (Attachment 4). The Resource Coordinator works with the consumer and family to complete sections on Functioning, Care Intensity and Organization, and Caregiver Needs and Strengths, while the Service Coordinator works with the consumer's provider(s) and, when appropriate, Department of Human Services caseworker to complete sections on Problem Presentation, Risk Behaviors, and Strengths. The Self-Efficacy Scale (Operational Manual, page 24) and an initial assessment of the consumer and family's Readiness for Change (Operational Manual, page 55) are also completed at this time. Completion of all measures is expected within one week of case assignment whenever possible, and results are entered into CACM. The Care Management dyad meets to review the collected data, and shares preliminary CANS-MH scores with the consumer, family, and provider. This consultation may result in adjustment of scores based on additional or new information provided. On the basis of the final CANS-MH scores, the Care Management dyad identifies CANS-MH areas with scores of 2 or 3 that require action, and begins to prioritize and establish Care Plan goals. All goals and interventions are developed in a collaborative process with the consumer and family, care providers as indicated, and the Care Management team. The Care Plan is entered into CACM and reviewed with ABC clinical management and a board-certified child psychiatrist.

Care Plans may be developed across four domains of functioning including: medical/clinical, social support, community involvement and self-management. The medical/clinical domain includes any goals related to the consumer's physical and behavioral health, as well as substance abuse issues. The social support domain includes goals related to the consumer and family's ability to engage in natural supports within the context of their family and community, including involvement in faith-based community activities. The community involvement domain includes goals related to any involved community agencies and more formalized support groups. This may include human service agencies such as the Department of Human Services, the school district, developmental disabilities and juvenile justice organizations, as well as participation in structured community support groups. The self-management domain includes goals related to the consumer and family's self-efficacy, need for illness education, development of advocacy skills, and overall engagement in the treatment process. A primary focus of all care planning is to ensure a family-focused approach to care that minimizes duplication of services and conflicting goals, and maximizes the potential for positive outcomes across the system of care. A core aim of AFFIRM Care Management is to provide consumers and families with the skills needed to take charge of their health care to promote the highest possible quality of life.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

G. Activity VII. Improvement Strategies. Real, sustained improvements in care result from a continuous cycle of measuring and analyzing performance, and developing and implementing system-wide improvements in care. Describe interventions designed to change behavior at an institutional, practitioner, or consumer level.

The Service Coordinator and Resource Coordinator work as a team to meet the goals developed throughout the care planning process, and coordinate their efforts on a regular basis in addition to weekly supervision meetings. Specific Care Management responsibilities of the Resource Coordinators may include the following activities:

- Weekly face-to-face or telephone contact with the consumer and family to assess their current status and assist the consumer and family in engaging in the treatment process as needed;
- Active coordination of treatment services (assisting the family in keeping appointments, assisting with necessary transportation, attending appointments if necessary to meet Care Plan goals);
- Review of physical health care needs such as accessing a PCP, and assisting in coordinating appropriate services;
- Active referral and coordination with necessary resources such as housing, food bank, clothing and utilities;
- Attendance and participation in school meetings such as Individual Educational Plan (IEP) meetings;
- Attendance and active participation at staffings including Team Decision Making meetings at the county Department of Human Services and treatment planning meetings at provider sites;
- AFFIRM Illness Self-Management Education, utilizing modules from Taking Charge…Choices for Better Health;
- Development of My Action Plan (MAP) or crisis plan;
- Quarterly assessment of the non-clinical scales on the CANS-MH;
- Ongoing assessment of the consumer and family's readiness for change, resource needs, barriers to care, and level of self-efficacy; and
- At least monthly review of Care Plan goals to assess appropriateness of goals and interventions, with an emphasis on the consumer and family's comfort with the goals and interventions, progress, and level of engagement in the Care Management process.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

G. Activity VII. Improvement Strategies. Real, sustained improvements in care result from a continuous cycle of measuring and analyzing performance, and developing and implementing system-wide improvements in care. Describe interventions designed to change behavior at an institutional, practitioner, or consumer level.

Service Coordinators engage in the following Care Management activities:

- Having no less than monthly (depending on level of service provided, e.g., daily for hospitalized consumers, bi-monthly for home-based services) contact with the provider regarding the current clinical status, progress toward goals, barriers to discharge;
- Having no less than monthly contact with the provider regarding interventions utilized and potential changes needed including mode of treatment, frequency of contacts, and medication management;
- Managing the authorization and reauthorization process;
- Ensuring that utilization management decisions are consistent with and support the Care Management plan;
- Consulting with the Associate Medical Director regarding areas of clinical concern as needed in addition to regular supervision;
- Facilitating provider access to the Associate Medical Director for diagnostic and medical management consultation as needed;
- Having no less than monthly contact with other involved agencies and service providers such as human services and the school district regarding coordination of care, joint treatment planning, and prioritizing issues of concern across the system of care when possible;
- Attending and participating in relevant staffings including Team Decision Making meetings at the county Department of Human Services and treatment planning meetings at provider sites;
- Assisting in accessing clinical services, including arranging for emergency intervention such as Mobile Crisis evaluations and respite care;
- Conducting quarterly assessment of the clinical scales on the CANS-MH; and
- Assisting in transition and discharge planning.

G. Activity VII. Improvement Strategies. Real, sustained improvements in care result from a continuous cycle of measuring and analyzing performance, and developing and implementing system-wide improvements in care. Describe interventions designed to change behavior at an institutional, practitioner, or consumer level.

AFFIRM Care Management activities are supervised through weekly meetings directed by the Associate Medical Director for behavioral health programs and co-facilitated by the clinical and advocacy management staff. All Care Management staff attend the weekly meetings to review cases enrolled in AFFIRM. Every case is reviewed at least monthly. Cases are also selected for review based on clinical acuity and/or need identified by the Care Coordination dyad or by the management staff. Cases and questions are presented for discussion and problem solving, recommendations are made, and goals established or revised. Information presented may include results of assessments, currently prescribed medications, the consumer's living situation, presenting problem and current progress toward goals, barriers to care and other relevant situational or systemic issues of concern. The discussion may lead to a decision that a consult by the Associate Medical Director with the provider is necessary to facilitate appropriate care. Ad hoc supervision by the clinical management team is also available as needed. In addition to supporting the Care Management dyad, supervision is utilized to monitor the appropriateness of goals, timeliness of goal attainment, barriers to care, and cost effectiveness of the Care Management process.

Remeasurement 1 to Remeasurement 2

The design of the AFFIRM intervention did not change for this measurement period. However, as discussed in section 10, staff encountered significant difficulty in meeting the turnaround times stated in the baseline to remeasurement 1 section. It should also be noted that Colorado Access underwent a significant downsizing and operational restructuring in 2006. While this did not directly impact the AFFIRM program, it did serve as a distraction to staff and resulted in a reduced focus on the program for a period of time this summer.

H. Activity VIIIa. Data analysis: Describe the data analysis process in accordance with the analysis plan and any adhoc analysis done on the selected clinical or nonclinical study indicators. Include the statistical analysis techniques utilized and *p* values.

Baseline Measurement (Indicators #1 and #2)

For indicator #1, inpatient readmission rates, two sets of measurements were obtained. First, for baseline measurement, a total of 33 enrolled consumers active in the AFFIRM program were identified. All inpatient claims for these consumers were extracted from the transaction system. We looked at admissions that occurred from January through the end of October 2004 and readmissions that occurred between January and December 31, 2004. This resulted in a total of 52 admissions to psychiatric inpatient facilities. Admissions were then examined by date of service. Of the 52 total admissions, there were 36 (69.2%) admissions that occurred prior to the consumer's enrollment in the program, and 2 readmissions within 30 days of the date of discharge from the prior admission, resulting in a readmission rate of 5.6%. Analysis was next conducted on a group of 44 consumers who were screened and risked-in, but not enrolled in or disenrolled from the program, constituting the comparison group. A transaction system extraction of all inpatient claims for these consumers dating a year prior to program start date was reviewed. Admissions from January through October 2004 were included as were readmissions from January to December 31, 2004. This analysis resulted in a total of 61 admissions to psychiatric inpatient facilities. Of these, there were 38 (62.3%) admissions that occurred prior to the consumer's risk-in date, and 9 readmissions within 30 days of the date of discharge from the prior admission, resulting in a readmission rate of 23.7%.

Review of full claims data drew attention to the number of emergency room visits and crisis contacts for both enrolled and non-enrolled groups, and led to the question of whether there might be differences between the groups. Given the high-risk nature of the population, an ad hoc analysis was performed for both groups on emergency room visits leading to an admission, and emergency room or crisis contacts not leading to an admission. In the enrolled consumer group, a total of 32 claims were found for emergency room visits leading to admission, and 11 claims were found for emergency room or crisis contacts not leading to admission. Thirteen consumers (39.4%) from the enrolled group had a total of 25 (78.1%) emergency room visits leading to admission prior to his or her enrollment in the program, and five consumers (15.2%) had a total of 7 (63.6%) emergency room or crisis contacts not leading to admission, prior to program enrollment. In the comparison group of non-enrolled consumers, a total of 33 claims were found for emergency room visits leading to admission, and 16 claims were found for emergency room or crisis contacts not leading to admission. Sixteen consumers (36.4%) from the non-enrolled group had a total of 23 (69.7%) emergency room visits leading to admission prior to his or her risk-in date, and nine consumers (20.5%) had a total of 11 (68.7%) emergency room or crisis contacts not leading to admission, prior to his or her risk-in date.

For indicator #2, reduction in CANS-MH domain scores, CANS-MH assessment scores were extracted from the CACM database, and all enrolled consumers who had an initial CANS-MH assessment upon enrollment and at least one follow-up CANS-MH assessment at quarterly intervals following enrollment were identified. Of the total number of enrolled consumers (thirty-three), there were fifteen consumers who had two or more completed CANS-MH assessments. Of the remainder, thirteen had an initial CANS-MH assessment but no follow-up CANS-MH assessment at

H. Activity VIIIa. Data analysis: Describe the data analysis process in accordance with the analysis plan and any adhoc analysis done on the selected clinical or nonclinical study indicators. Include the statistical analysis techniques utilized and *p* values.

the time of this report, and five did not have a completed initial CANS-MH assessment primarily due to program start date late in the year. For the fifteen enrolled consumers with two or more completed CANS-MH assessments, initial CANS-MH item scores for each of the six CANS-MH domains were summed to yield domain scores. These initial CANS-MH domain scores establish the baseline measure.

For indicator #2, reduction in CANS-MH domain scores, remeasurement consisted of extracting from the CACM database all follow-up CANS-MH assessment scores that were obtained at three months, and if available, six months following enrollment in the program. For the fifteen enrolled consumers with two or more completed CANS-MH assessments, item scores for each of the six CANS-MH domains were summed to yield domain scores for the three-month and six-month assessment periods. The most recent CANS-MH domain scores were then compared to the initial CANS-MH domain scores to determine whether there had been a reduction of one point or more in any of the six domains. Of the fifteen enrolled consumers with an initial and at least one follow-up CANS-MH assessment, ten (66.7%) demonstrated a reduction of at least one point in four of six domains. Of these, five (33.3%) showed improvement in four of six domains, two (13.3%) showed improvement in five of six domains, and three (20.0%) showed improvement in all six domains. Of the remainder, one (6.7%) showed improvement in one of six domains, three (20.0%) showed improvement in two of six domains, and one (6.7%) showed improvement in three of six domains. Reductions in domain scores from baseline to remeasurement assessments ranged from one to eight points.

Ad hoc analysis of scores by domain indicated reductions in each of the six domains in total points and average points calculated across all fifteen consumers. In the Problem Presentation domain, ten (66.7%) consumers obtained lower scores upon remeasurement, 2 (13.3%) showed no change in score from baseline to remeasurement, and 3 (20.0%) obtained higher scores upon remeasurement. In the Risk Behaviors domain, ten (66.7%) consumers obtained lower scores upon remeasurement, 3 (20.0%) showed no change from baseline to remeasurement, and 2 (13.3%) obtained higher scores upon remeasurement. In the Functioning domain, eight (53.3%) consumers obtained lower scores upon remeasurement, 2 (13.3%) showed no change from baseline to remeasurement, and 5 (33.3%) obtained higher scores upon remeasurement. In the Care Intensity and Organization domain, ten (66.7%) consumers obtained lower scores upon remeasurement, 2 (13.3%) showed no change from baseline to remeasurement, and 3 (20.0%) obtained higher scores upon remeasurement. In the Caregiver Needs and Strengths domain, ten (66.7%) consumers obtained lower scores upon remeasurement, 0 (0.0%) showed no change from baseline to remeasurement, and 5 (33.3%) obtained higher scores upon remeasurement. Lastly, in the Strengths domain, ten (66.7%) consumers obtained lower scores upon remeasurement, 1 (6.7%) showed no change from baseline to remeasurement, and 4 (26.7%) obtained higher scores upon remeasurement.

Results by consumer and by domain can be found in Attachment 7.

No statistical analysis is possible for this reporting period as only one set of measurements of CANS-MH domain scores are obtained comparing baseline to remeasurement.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

H. Activity VIIIa. Data analysis: Describe the data analysis process in accordance with the analysis plan and any adhoc analysis done on the selected clinical or nonclinical study indicators. Include the statistical analysis techniques utilized and *p* values.

Remeasurement 1 (Indicator #1)

For indicator #1, inpatient readmission rates, remeasurement was conducted in February 2005 on all inpatient claims extracted from the transaction system for enrolled consumers. Admissions were examined by date of service. Of the total 52 admissions to psychiatric inpatient facilities obtained for the 33 consumers in the enrolled group, there were 16 (30.8%) admissions that occurred after the consumer's date of enrollment in the program, and 3 readmissions within 30 days of the date of discharge from the prior admission, resulting in a readmission rate of 18.8%. A t-test calculating variance ranges of the percentages and determining the degree of overlap at the 95% confidence level indicated that the difference between readmission rates pre-enrollment (baseline) and post-enrollment (remeasurement) was not statistically significant (p=0.558).

Analysis was next conducted on the comparison group of 44 consumers who were screened and risked-in, but not enrolled in or disenrolled from the program. Of the total 61 admissions to psychiatric inpatient facilities obtained for this group of non-enrolled consumers, there were 23 (37.7%) admissions that occurred after the consumer's risk-in date and 4 readmissions within 30 days of the date of discharge from the prior admission, resulting in a readmission rate of 17.4%.

The ad hoc analysis that was performed for both groups on emergency room visits leading to an admission, and emergency room or crisis contacts not leading to an admission, included identification of emergency room or crisis contacts occurring after a consumer's enrollment or risk-in date, to determine if there was a difference between enrolled and non-enrolled groups that might be related to program outcomes. In the enrolled consumer group, of the 32 claims found for emergency room visits leading to admission and 11 claims found for emergency room or crisis contacts not leading to admission, seven consumers (21.2%) had a total of 7 (21.9%) emergency room visits leading to admission after the date of his or her enrollment in the program, and four consumers (6.1%) had a total of 4 (36.4%) emergency room or crisis contacts not leading to admission, after the date of program enrollment. In the comparison group of non-enrolled consumers, of the 33 claims found for emergency room visits leading to admission and 16 claims found for emergency room or crisis contacts not leading to admission, seven consumers (15.9%) had a total of 9 (30.3%) emergency room visits leading to admission after his or her risk-in date, and five consumers (11.4%) had a total of 5 (31.3%) emergency room or crisis contacts not leading to admission, following his or her risk-in date.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

H. Activity VIIIa. Data analysis: Describe the data analysis process in accordance with the analysis plan and any adhoc analysis done on the selected clinical or nonclinical study indicators. Include the statistical analysis techniques utilized and *p* values.

Remeasurement 1 (Indicator #2)

For indicator #2, reduction in three or six month CANS-MH domain scores, remeasurement was captured based on any baseline CANS-MH beginning on 4/1/06 and any follow-up CANS-MH measures through 2/1/07. Of the 5 members enrolled in the program, 1 had a reduction in at least four of the six CANS-MH domains, resulting in a reduction rate of 20.0%. This change was not statistically significant (p=0.202).

Remeasurement 2 (Indicator #1)

For indicator #1, inpatient readmission rates, remeasurement was conducted in February 2007 on all inpatient claims extracted from the transaction system for enrolled members who were admitted between April 1, 2006 and October 31, 2006. Of the 5 members enrolled in the AFFIRM program, there were a total of 5 admissions during the measurement period (2 members accounted for the 5 admits). None of the members admitted to the hospital were readmitted within 30 days, resulting in a readmission rate of 0%. The change in readmission rates from baseline (5.56%) to remeasurement 1 was not statistically significant (18.75%) (p=0.558), but the change from baseline (5.56%) to remeasurement 2 (0.0%) was statistically significant (p=0.002).

Remeasurement 3

N/A

H. Activity VIIIb. Interpretation of study results: Describe the results of the statistical analysis, interpret the findings, and discuss the successfulness of the study and indicate follow-up activities. Also, identify any factors that could influence the measurement or validity of the findings.

Baseline Measurement

For indicator #1, inpatient readmission rates, two sets of measures were obtained. For baseline measurement of enrolled consumers whose claims for admission to a psychiatric inpatient facility were analyzed for the period prior to his or her date of enrollment in AFFIRM, there were only two readmissions within 30 days of discharge from a previous admission, of a total 36 admissions, for a readmission rate of 5.6%. For the comparison group of consumers were screened and risked-in but not enrolled in the program, there were nine readmissions within 30 days of discharge from a previous admission, of a total 38 admissions, for a readmission rate of 23.7%. While the difference in readmission rates between the two groups was not statistically significant (p=0.232), it may suggest possible qualitative dissimilarities between the groups. For example, because many in the non-enrolled group were eligible for participation in AFFIRM but declined involvement or withdrew involvement, these families may feel already encumbered by the obligations resulting from their involvement with multiple systems and unable or unwilling to commit themselves to an additional program. Alternately, they may be more likely to underestimate the extent or severity of their child's problems and believe that further attention or resources are not needed.

The ad hoc analysis of emergency room visits and crisis contacts for both the enrolled and non-enrolled groups to determine possible differences between the groups indicates that the enrolled group had a higher percentage of emergency room visits leading to an admission (78.1%) prior to enrollment date, than the non-enrolled group (69.7%) did prior to risk-in date. Percentages of emergency room or crisis contacts not leading to an admission, prior to enrollment or risk-in date, were not markedly different between the two groups (63.6% for enrolled consumers versus 68.7% for non-enrolled consumers).

For indicator #2, baseline measurement of CANS-MH domain scores indicate only the severity of the consumer's status upon enrollment in the AFFIRM program. Clinically, CANS-MH scores for each consumer drive the development of Care Plans and the implementation of intervention strategies specific to the consumer's and family's needs as indicated by his or her scores.

For indicator #2, the goal of 60% of consumers ages 0-17 enrolled in the AFFIRM program with three-month or six-month follow-up CANS-MH assessments, as determined by length of enrollment, showing a reduction of one point or more in four of six domain scores was exceeded. Follow-up CANS-MH domain scores were lower than at least one point in four of six domains for ten of fifteen consumers, or 66.7%. This reduction in CANS-MH domain scores demonstrates that AFFIRM Care Management activities conducted by Service Coordinator and Resource Coordinator staff have been successful to date in improving outcomes for enrolled consumers and families. Ratings for the domains suggest improved clinical and functional status, increased consumer safety, enhanced consumer and caregiver skills and strengths, and greater stability in treatment and home environments. This inference is strengthened by the fact that 74% of the cases in which a lower domain score was obtained

H. Activity VIIIb. Interpretation of study results: Describe the results of the statistical analysis, interpret the findings, and discuss the successfulness of the study and indicate follow-up activities. Also, identify any factors that could influence the measurement or validity of the findings.

upon remeasurement showed a reduction of at least two points, and 59% showed a reduction of at least three points, indicating greater degrees of improvement and exceeding the expectations defined in the goal. Additionally, half of the consumers showing progress as evidenced by reductions in domain scores obtained improvements in five or six of the domains, rather than four, also exceeding the expectations defined in the goal.

In the aggregate, of the 90 total CANS-MH domain scores (6 domains x 15 consumers), over 64% showed a reduction. These gains are further reflected in the overall reductions in total points and average points in each of the domains across all consumers. The Problem Presentation domain showed a 9% reduction, the Risk Behaviors domain showed a 21% reduction, the Functioning domain showed a 14% reduction, the Care Intensity domain showed a 21% reduction, the Caregivers Needs and Strengths domain showed a 16% reduction, and the Strengths domain showed a 5% reduction. In each domain, over half to two-thirds of the enrolled consumers obtained scores indicating improvement. The fact that one-third to half of the consumers showed no change or worsening status suggests that there was no rating bias or threats to the validity of the assessments.

For those enrolled consumers for whom both utilization and CANS-MH remeasurement data is available, some findings of interest were observed. One enrolled consumer with a readmission within 30 days following date of enrollment obtained higher CANS-MH scores (poorer results) on remeasurement of the Problem Presentation domain, indicating increased symptoms, and Risk Behaviors. Another consumer who had an inpatient readmission within 30 days prior to enrollment but no readmissions within 30 days after enrollment obtained improved CANS-MH scores in four of six domains upon remeasurement. Additionally, six enrolled consumers with follow-up CANS-MH assessments that obtained improved CANS-MH scores in at least one domain also showed a reduction in emergency room visits leading to an inpatient admission. These findings, though preliminary, tend to support the effectiveness of the program. Relationships between utilization patterns and CANS-MH scores will be further examined in future reports as enrollment numbers increase and trends may be identified.

While the numbers for this initial reporting period are small, Access Behavioral Care is encouraged by the positive results reflected in lower inpatient, emergency room and crisis utilization, and reduced CANS-MH scores, for this high-risk population. We believe that this intervention strategy is of broad benefit to members and that additional time and experience with the program will confirm observed trends in improved consumer and family outcomes. The AFFIRM Care Management program, versus “usual care”, provides opportunities for more comprehensive treatment plans based on information obtained from CANS-MH assessments completed in collaboration with the consumer, family, and provider, a more structured approach to intervention, and a higher level of supervision and oversight. Strategies will be refined as needed on the basis of lessons learned in staff’s experience with this most challenging and complex group of consumers and families.

H. Activity VIIIb. Interpretation of study results: Describe the results of the statistical analysis, interpret the findings, and discuss the successfulness of the study and indicate follow-up activities. Also, identify any factors that could influence the measurement or validity of the findings.

Remeasurement 1 (indicator #1)

For indicator #1, inpatient readmission rates, the goal of a 30-day readmission rate for enrolled consumers less than or equal to the Children's Mental Health Benchmarking Project mean of 12.0% was not met. However, the difference between the obtained 30-day readmission rate of 18.8% for enrolled consumers following their date of enrollment, when compared to the 5.6% readmission rate prior to enrollment, was not statistically significant (p=0.121). Figures for the first reporting period are based on very small numbers, and it is likely too soon to determine program impacts on readmission rates. The 30-day readmission rate for non-enrolled consumers following their risk-in date is similar at 17.4% and suggests little difference between the enrolled and non-enrolled groups at this point in the project. It is anticipated that greater impact on inpatient readmission rates for enrolled consumers will be seen as length of AFFIRM program enrollment and participation increases, and Care Management staff, families and providers collaborate to achieve higher levels of stability and treatment at lower levels of care.

Other than 30-day readmission rates, there were indications of potential program effects observed from baseline to remeasurement. For enrolled consumers, the rate of all admissions declined from 69.2% prior to enrollment date, to 30.8% after the enrollment date. This compares to 62.3% prior to risk-in date, to 37.7% after the risk-in date, for the non-enrolled group. Similarly, for the enrolled group, the percentage of emergency room visits leading to an admission declined from 78.1% prior to enrollment date, to 21.9% after the enrollment date. For the non-enrolled group, the percentages for emergency room visits leading to admission were 69.7% prior to risk-in date, and 30.3% after the risk-in date.

For all emergency room and crisis contacts (both with inpatient admission and without inpatient admission), the rate of emergency contact declined from 74.4% prior to enrollment date, to 25.6% after the enrollment date, for enrolled consumers. This compares to rates of emergency contact of 69.4% prior to risk-in date, and 30.6% after the risk-in date, for non-enrolled consumers. Although none of the differences between the enrolled and non-enrolled groups was found to be statistically significant, it does appear that enrollment in the AFFIRM program has had some initial positive effects on reducing overall inpatient and emergency utilization. These trends are promising and will continue to be monitored.

Findings related to utilization patterns are considered valid, although possibly limited by the fact that this is the first year of implementation of the program. Consequently, at this point, with fairly small enrollment numbers, definitive conclusions cannot yet be drawn about significant change in inpatient utilization. Remeasurement in future reporting periods will span more time and provide a larger claims data set for analysis of inpatient utilization, to determine whether the currently observed trends hold.

H. Activity VIIIb. Interpretation of study results: Describe the results of the statistical analysis, interpret the findings, and discuss the successfulness of the study and indicate follow-up activities. Also, identify any factors that could influence the measurement or validity of the findings.

Remeasurement 1 (Indicator #2)

During the baseline measurement period, 66.7% of members demonstrated reductions in CANS-MH scores compared to 20.0% of members during remeasurement 1. Although this is a substantial decline, it does not represent a statistically significant reduction (p=0.202). Interpretation of the results for this indicator is difficult due to the small number of members enrolled in the program during remeasurement 1.

Remeasurement 2 (Indicator #1)

Although a small number of members were enrolled in the AFFIRM program during remeasurement 2 for indicator #1, no readmissions based on five inpatient admits is a positive outcome. Even based on only five admissions, the change from baseline to remeasurement 1 was statistically significant (p=0.002).

Remeasurement 3

N/A

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

I. Activity IX. Study Results Summary and Improvement: List study results and describe any meaningful change in performance observed during the time period of analysis.

#1 Quantifiable Measure: Psychiatric inpatient readmission rates at 30 days

Time Period Measurement Covers	Baseline Project Indicator Measurement	Numerator	Denominator	Rate or Results	Industry Benchmark	Statistical Test and Significance*
04/01/04 - 02/01/05	***Baseline:***	2	36	5.6%	12.0%	N/A
04/01/05 - 02/01/06	**Remeasurement 1:**	3	16	18.8%	12.0%	B-1, p=0.558
04/01/06 - 02/01/07	**Remeasurement 2:**	0	5	0%	12.0%	1-2, p=0.532; B-2 p=0.002*
	Remeasurement 3:					
	Remeasurement 4:					
	Remeasurement 5:					

#2 Quantifiable Measure: Reduction of three- or six-month CANS-MH domain scores from initial CANS-MH domain scores

Time Period Measurement Covers	Baseline Project Indicator Measurement	Numerator	Denominator	Rate or Results	Industry Benchmark	Statistical Test and Significance*
04/01/05 - 02/01/06	***Baseline:***	10	15	66.7%	N/A	N/A
04/01/06 - 02/01/07	**Remeasurement 1:**	1	5	20.0%	N/A	B-1, p=0.202
	Remeasurement 2:	N/A	N/A	N/A	N/A	N/A
	Remeasurement 3:					
	Remeasurement 4:					
	Remeasurement 5:					

- Indicates a statistically significant difference. See Attachment 8 for details of analysis.

Appendix A: PIP Summary Form: Improving Outcomes for High-Risk Youth Through AFFIRM Care Management *for* Access Behavioral Care

J. Activity X. Sustained improvement: Describe any demonstrated improvement through repeated measurements over comparable time periods. Discuss any random year-to-year variation, population changes, and sampling error that may have occurred during the remeasurement process.

Sustained improvement in readmission rates and outcomes from the AFFIRM program has not been demonstrated. Although improvement in readmission rates from baseline to remeasurement 2 was statistically significant, the difference was clinically negligible. The outcomes based on CANS-MH scores showed that 20% of members enrolled during remeasurement 1 made improvements compared to 66.7% in the baseline period. Although the decline based on the CASNS-MH is large (46.7%), the small number of members enrolled in this measurement period prevents any conclusive judgments about the success of the intervention.

What does appear more conclusive is the ongoing feasibility of the AFFIRM program. In the program's initial year only 33 members were enrolled. The low enrollment was judged to be due to the newness of the program. Enrollment was anticipated to increase in the following year. However, enrollment dramatically decreased in the following year. We conducted a significant review of the AFFIRM program, the data collection expectations, the value of the data collected and the overall staffing model of the program.

Our review determined that the program as designed required significantly more resources than originally anticipated. The collection of data to complete the entire CANS assessment required multiple meetings with the family and the adolescent, receipt and review of the adolescent's entire medical record and the clinical diagnosis of the adolescent by internal health plan staff as well as meetings with those in the educational system and other professionals involved with the adolescent's care. Frequently, the initial CANS assessment was not completed prior to the initiation of the three month CANS assessment. AFFIRM staff also reported that they were spending all their time collecting CANS data and had no time to administer any of the program's interventions.

A review of the staffing model of the program completed by the management team revealed that internal health plan staff were duplicating activities the adolescent's clinical team was responsible for. In addition, the AFFIRM program as designed did not have dedicated staff. The program was to be administered by the behavioral health utilization management staff. This staff is responsible for the review and coordination of all services for the entire BHO. This is a significant day to day responsibility and without specific time carved out for the AFFIRM program, the activities associated with the program were usurped by utilization management responsibilities. The care management programs in the physical health lines of business have assigned staff whose primary responsibility is the screening, enrollment and management of members in the care management program.

A redesign of the AFFIRM program resulted in the assignment of a single resource dedicated solely to the AFFIRM program. This is consistent with the design and an operation of the company's other successful care management program.

J. Activity X. Sustained improvement: Describe any demonstrated improvement through repeated measurements over comparable time periods. Discuss any random year-to-year variation, population changes, and sampling error that may have occurred during the remeasurement process.

The review of the failure to achieve sustained improvement in indicator 1, readmission rates, is attributed to the small number of members enrolled in the program. While reduction in readmission scores is an admirable goal, with such a small number of members enrolled in the program and the fact that the baseline readmission rate was 5.6%, we would have had to reduce readmissions to 0 to show any significant improvement. The fact that the baseline readmission rate was already below the benchmark of 12% suggests that this may not have been the best indicator to select to measure the impact of this program. While it was anticipated that the AFFIRM program would be effective in reducing readmissions, the implementation of the program as designed did not allow for enough enrollment into the program to detect significant changes in this rate.

As a result of the deficiencies in the design of the AFFIRM intervention and the outcomes being measured, Colorado Access management has determined it is no longer possible to continue this program as originally designed. After our review of the amount of resources dedicated to the program and our failure to demonstrate significant improvement in either of our identified measures, we feel that refocusing the resources now devoted to AFFIRM to higher priority areas that serve a greater number of members is the most appropriate resolution to these issues.

In a technical assistance call with HCPF and HSAG, these issues were discussed and it was determined that the best course of action was to initiate a new PIP where the available resources can be better targeted to achieve positive improvements for our members.

We look forward to discussion with HSAG around the development of a new Performance Improvement Project that is better aligned with the needs of our members and focused on a high priority area of care where we see needed improvement in the quality of mental health care delivered to our members.

PIPs provide a structured method of assessing and improving the processes, and thereby outcomes, of care for the population that a BHO serves. This structure facilitates the documentation and evaluation of improvements in care or service. PIPs are conducted by the BHOs to assess and improve the quality of clinical and nonclinical health care services received by consumers.

The PIP evaluation is based on CMS guidelines as outlined in the CMS publication, *Validating Performance Improvement Projects, A Protocol for Use in Conducting Medicaid External Quality Review Activities*, Final Protocol, Version 1.0, May 1, 2002 (CMS PIP Protocol).

This document highlights the rationale for each activity as established by CMS. The protocols for conducting PIPs can be used to assist the BHOs in complying with requirements.

CMS Rationale

Activity I. Appropriate Study Topic

All PIPs should target improvement in relevant areas of clinical care and nonclinical services. Topics selected for study by Medicaid managed care organizations must reflect the BHO's Medicaid enrollment in terms of demographic characteristics, prevalence of disease, and the potential consequences (risks) of disease (CMS PIP Protocol, page 2).

Activity II. Clearly Defined, Answerable Study Question

It is important for the BHO to clearly state, in writing, the question(s) the study is designed to answer. Stating the question(s) helps maintain the focus of the PIP and sets the framework for data collection, analysis, and interpretation (CMS PIP Protocol, page 5).

Activity III. Clearly Defined Study Indicator(s)

A study indicator is a quantitative or qualitative characteristic (variable) reflecting a discrete event (e.g., an older adult has/has not received an influenza vaccination in the last 12 months) or a status (e.g., a consumer's blood pressure is/is not below a specified level) that is to be measured.

Each project should have one or more quality indicators for use in tracking performance and improvement over time. All indicators must be objective, clearly and unambiguously defined, and based on current clinical knowledge or health services research. In addition, all indicators must be capable of objectively measuring either consumer outcomes, such as health status, functional status, or consumer satisfaction, or valid proxies of these outcomes.

Indicators can be few and simple, many and complex, or any combination thereof, depending on the study question(s), the complexity of existing practice guidelines for a clinical condition, and the availability of data and resources to gather the data.

Indicator criteria are the set of rules by which the data collector or reviewer determines whether an indicator has been met. Pilot or field testing is helpful in the development of effective indicator criteria. Such testing allows the opportunity to add criteria that might not have been anticipated in the design phase. In addition, criteria are often refined over time based on results of previous studies. However, if criteria are changed significantly, the method for calculating an indicator will not be consistent and performance on indicators will not be comparable over time.

It is important, therefore, for indicator criteria to be developed as fully as possible during the design and field testing of data collection instruments (CMS PIP Protocol, page 5).

Activity IV. Use a Representative and Generalizable Study Population

Once a topic has been selected, measurement and improvement efforts must be systemwide (i.e., each project must represent the entire Medicaid enrolled population to which the PIP study indicators apply). Once that population is identified, the BHO must decide whether to review data for that entire population or use a sample of that population. Sampling is acceptable as long as the samples are representative of the identified population (CMS PIP Protocol, page 8). (See "Activity V. Valid Sampling Techniques.")

Activity V. Valid Sampling Techniques

If the BHO uses a sample to select consumers for the study, proper sampling techniques are necessary to provide valid and reliable (and therefore generalizable) information on the quality of care provided. When conducting a study designed to estimate the rates at which certain events occur, the sample size has a large impact on the level of statistical confidence in the study estimates. Statistical confidence is a numerical statement of the probable degree of certainty or accuracy of an estimate. In some situations, it expresses the probability that a difference could be due to chance alone. In other applications, it expresses the probability of the accuracy of the estimate. For example, a study may report that a disease is estimated to be present in 35 percent of the population. This estimate might have a 95 percent level of confidence, plus or minus 5 percentage points, implying a 95 percent certainty that between 30 percent and 40 percent of the population has the disease.

The true prevalence or incidence rate for the event in the population may not be known the first time a topic is studied. In such situations, the most prudent course of action is to assume that a maximum sample size is needed to establish a statistically valid baseline for the project indicators (CMS PIP Protocol, page 9).

Activity VI. Accurate/Complete Data Collection

Procedures used by the BHO to collect data for its PIP must ensure that the data collected on the PIP indicators are valid and reliable. Validity is an indication of the accuracy of the information obtained. Reliability is an indication of the repeatability or reproducibility of a measurement. The BHO should employ a data collection plan that includes:

- Clear identification of the data to be collected.
- Identification of the data sources and how and when the baseline and repeat indicator data will be collected.
- Specification of who will collect the data.
- Identification of instruments used to collect the data.

When data are collected from automated data systems, development of specifications for automated retrieval of the data should be devised. When data are obtained from visual inspection of medical records or other primary source documents, several steps should be taken to ensure the data are consistently extracted and recorded:

1. The key to successful manual data collection is in the selection of the data collection staff. Appropriately qualified personnel, with conceptual and organizational skills, should be used to abstract the data. However, their specific skills should vary depending on the nature of the data collected and the degree of professional judgment required. For example, if data collection involves searching throughout the medical record to find and abstract information or judge whether clinical criteria were met, experienced clinical staff, such as registered nurses, should collect the data. However, if the abstraction involves verifying the presence of a diagnostic test report, trained medical assistants or medical records clerks may be used.
2. Clear guidelines for obtaining and recording data should be established, especially if multiple reviewers are used to perform this activity. The BHO should determine the necessary qualifications of the data collection staff before finalizing the data collection instrument. An abstractor would need fewer clinical skills if the data elements within the data source are more clearly defined. Defining a glossary of terms for each project should be part of the training of abstractors to ensure consistent interpretation among project staff.
3. The number of data collection staff used for a given project affects the reliability of the data. A smaller number of staff members promotes interrater reliability; however, it may also increase the amount of time it takes to complete this task. Intrarater reliability (i.e., reproducibility of judgments by the same abstractor at a different time) should also be considered (CMS PIP Protocol, page 12).

Activity VII. Appropriate Improvement Strategies

Real, sustained improvements in care result from a continuous cycle of measuring and analyzing performance and developing and implementing systemwide improvements in care. Actual improvements in care depend far more on thorough analysis and implementation of appropriate solutions than on any other steps in the process.

An improvement strategy is defined as an intervention designed to change behavior at an institutional, practitioner, or consumer level. The effectiveness of the intervention activity or activities can be determined by measuring the BHO's change in performance, according to predefined quality indicators. Interventions are key to an improvement project's ability to bring about improved health care outcomes. Appropriate interventions must be identified and/or developed for each PIP to ensure the likelihood of causing measurable change.

If repeat measures of quality improvement (QI) indicate that QI actions were not successful (i.e., the QI actions did not achieve significant improvement), the problem-solving process begins again with data analysis to identify possible causes, propose and implement solutions, and so forth. If QI actions were successful, the new processes should be standardized and monitored (CMS PIP Protocol, page 16).

Activity VIII. Sufficient Data Analysis and Interpretation

Review of the BHO data analysis begins with examining the BHO's calculated plan performance on the selected clinical or nonclinical indicators. The review examines the appropriateness of, and the BHO's adherence to, the statistical analysis techniques defined in the data analysis plan (CMS PIP Protocol, page 17).

Activity IX. Real Improvement Achieved

When a BHO reports a change in its performance, it is important to know whether the reported change represents real change, is an artifact of a short-term event unrelated to the intervention, or is due to random chance. The external quality review organization (EQRO) will need to assess the probability that reported improvement is actually true improvement. This probability can be assessed in several ways, but is most confidently assessed by calculating the degree to which an intervention is statistically significant. While this protocol does not specify a level of statistical significance that must be met, it does require that EQROs assess the extent to which any changes in performance reported by a BHO can be found to be statistically significant. States may choose to establish their own numerical thresholds for finding reported improvements to be significant (CMS PIP Protocol, page 18).

Activity X. Sustained Improvement Achieved

Real change results from changes in the fundamental processes of health care delivery. Such changes should result in sustained improvements. In contrast, a spurious, one-time improvement can result from unplanned accidental occurrences or random chance. If real change has occurred, the BHO should be able to document sustained improvement (CMS PIP Protocol, page 19).

Appendix C. Definitions and Explanations by Activity for Access Behavioral Care

This document was developed by HSAG as a resource to assist BHOs in understanding the broad concepts in each activity related to PIPs. The specific concept is delineated in the left column, and the explanations and examples are provided in the right column.

Definitions and Explanations	
Activity I. Appropriate Study Topic	
Broad Spectrum of Care	◆ Clinical focus areas: includes prevention and care of acute and chronic conditions and high volume/high-risk services. High-risk procedures may also be targeted (e.g., care received from specialized centers). ◆ Nonclinical areas: continuity or coordination of care addressed in a manner in which care is provided from multiple providers and across multiple episodes of care (e.g., disease-specific or condition-specific care).
Eligible Population	◆ May be defined as consumers who meet the study topic parameters.
Selected by the State	◆ If the study topic was selected by the state Medicaid agency, this information is included as part of the description under Activity One: Choose the Selected Study Topic in the PIP tool.
Activity II. Clearly Defined, Answerable Study Question	
Study Question	◆ The question(s) directs and maintains the focus of the PIP and sets the framework for data collection, analysis, and interpretation. The question(s) must be measurable and clearly defined. ◆ Examples: 1. Does outreach immunization education increase the rates of immunizations for children 0–2 years of age? 2. Does increasing flu immunizations for consumers with chronic asthma impact overall health status? 3. Will increased planning and attention to follow-up after inpatient discharge improve the rate of mental health follow-up services?

Definitions and Explanations	
Activity III. Clearly Defined Study Indicator(s)	
Study Indicator	◆ A quantitative or qualitative characteristic reflecting a discrete event or status that is to be measured. Indicators are used to track performance and improvement over time. ◆ Example: The percentage of enrolled consumers who were 12–21 years of age who had at least one comprehensive well-care visit with a primary care practitioner or an obstetrician-gynecologist during the measurement year.
Sources Identified	◆ Documentation/background information that supports the rationale for the study topic, study question, and indicators. ◆ Examples: HEDIS measures, medical community practice guidelines, evidence-based practices, or provider agreements. ◆ Practice guideline examples: American Academy of Pediatrics and American Diabetes Association.
Activity IV. Use a Representative and Generalizable Study Population	
Eligible Population	◆ Refers to consumers who are included in the study. ◆ Includes age, conditions, enrollment criteria, and measurement periods. ◆ Example: the eligible population includes all children ages 0–2 as of December 31 of the measurement period, with continuous enrollment and no more than one enrollment gap of 30 days or less.
Activity V. Valid Sampling Techniques	
True or Estimated Frequency of Occurrence	◆ This may not be known the first time a topic is studied. In this case, assume that a maximum sample size is needed to establish a statistically valid baseline for the study. HSAG will review whether the BHOs defined the impact the topic has on the population or the number of eligible consumers in the population.
Sample Size	◆ Indicates the size of the sample to be used.
Representative Sample	◆ Refers to the sample resembling the entire population.
Confidence Level	◆ Statistical confidence is a numerical statement of the probable degree of certainty or accuracy of an estimate (e.g., 95 percent level of confidence with a 5 percent margin of error).

Definitions and Explanations	
Activity VI. Accurate/Complete Data Collection	
Data Elements	◆ Identification of data elements includes unambiguous definitions of data that will be collected (e.g., the numerator/denominator, laboratory values).
Interrater Reliability (IRR)	◆ The HSAG review team evaluates if there is a tool, policy, and/or process in place to verify the accuracy of the data abstracted. Is there an over-read (IRR) process of a minimum-percentage review? ◆ Examples: a policy that includes how IRR is tested, documentation of training, and instruments and tools used.
Algorithms	◆ The development of any systematic process that consists of an ordered sequence of steps. Each step depends on the outcome of the previous step. ◆ The HSAG review team looks for the BHOs to describe the process used in data collection. What are the criteria (e.g., what Current Procedural Terminology and/or source codes were used)?
Data Completeness	◆ For the purposes of PIP scoring, data completeness refers to the degree of complete administrative data (e.g., encounter data or claims data). BHOs that compensate their providers on a fee-for-service basis require a submission of claims for reimbursement. However, providers generally have several months before they must submit the claim for reimbursement, and processing claims by the health plan may take several additional months, creating a claims lag. Providers paid on a capitated or salaried basis do not need to submit a claim to be paid, but should provide encounter data for the visit. In this type of arrangement, some encounter data may not be submitted. ◆ PIPs that use administrative data need to ensure the data has a high degree of data completeness prior to its use. Evidence of data completeness levels may include claim processing lag reports, trending of provider submission rates, policies and procedures regarding timeliness requirements for claims and encounter data submission, encounter data submission studies, and comparison reports of claims/encounter data versus medical record review. Discussion in the PIP should focus on evidence at the time the data was collected for use in identifying the population, sampling and/or calculation of the study indicators. Statements such as, "Data completeness at the time of the data pull was estimated to be 97.8 percent based on claims lag reports (see attached Incurred But Not Reported report)," along with the attachment mentioned, usually (but not always) are sufficient evidence to demonstrate data completeness.

Definitions and Explanations	
Activity VII. Appropriate Improvement Strategies	
Causes and Barriers	◆ Interventions for improvement are identified through evaluation or barrier analysis. If there was no improvement, what problem-solving processes were put in place to identify possible causes and proposed changes to implement solutions? ◆ It is expected that interventions associated with improvement of quality indicators will be system interventions.
Standardized	◆ If the interventions have resulted in successful outcomes, the interventions should continue and the BHO should monitor to assure the outcomes remain. ◆ Examples: if an intervention is the use of practice guidelines, then the BHOs continue to use them; if mailers are a successful intervention, then the BHOs continue the mailings and monitor outcomes.
Activity VIII. Sufficient Data Analysis and Interpretation	
Analysis Plan	◆ Each study should have a plan for how data analysis will occur. ◆ The HSAG review team will ensure that this plan was followed.
Generalization to the Study Population	◆ Study results can be applied to the general population with the premise that comparable results will occur.
Factors that Threaten Internal and External Validity	◆ Did the analysis identify any factors (internal or external) that would threaten the validity of study results? ◆ Example: there was a change in record extraction (e.g., a vendor was hired or there were changes in HEDIS methodology).
Presentation of the Data Analysis	◆ Results should be presented in tables or graphs with measurement periods, results, and benchmarks clearly identified.
Identification of Initial Measurement and Remeasurement of Study Indicators	◆ Clearly identify in the report which measurement period the indicator results reflect.
Statistical Differences Between Initial Measurement and Remeasurement Periods	◆ The HSAG review team looks for evidence of a statistical test (e.g., a t-test, or chi square test).
Identification of the Extent to Which the Study Was Successful	◆ The HSAG review team looks for improvement over several measurement periods. ◆ Both interpretation and analysis should be based on continuous improvement philosophies such that the BHO document data results and what follow-up steps will be taken for improvement.

Definitions and Explanations	
Activity IX. Real Improvement Achieved	
Remeasurement Methodology Is the Same as Baseline	◆ The HSAG review team looks to see that the study methodology remained the same for the entire study.
Documented Improvement in Processes or Outcomes of Care	◆ The study report should document how interventions were successful in impacting system processes or outcomes. ◆ Examples: there was a change in data collection or a rate increase or decrease demonstrated in graphs/tables.
Activity X. Sustained Improvement Achieved	
Sustained Improvement	◆ The HSAG review team looks to see if study improvements have been sustained over the course of the study. This needs to be demonstrated over a period of several (more than two) remeasurement periods.