

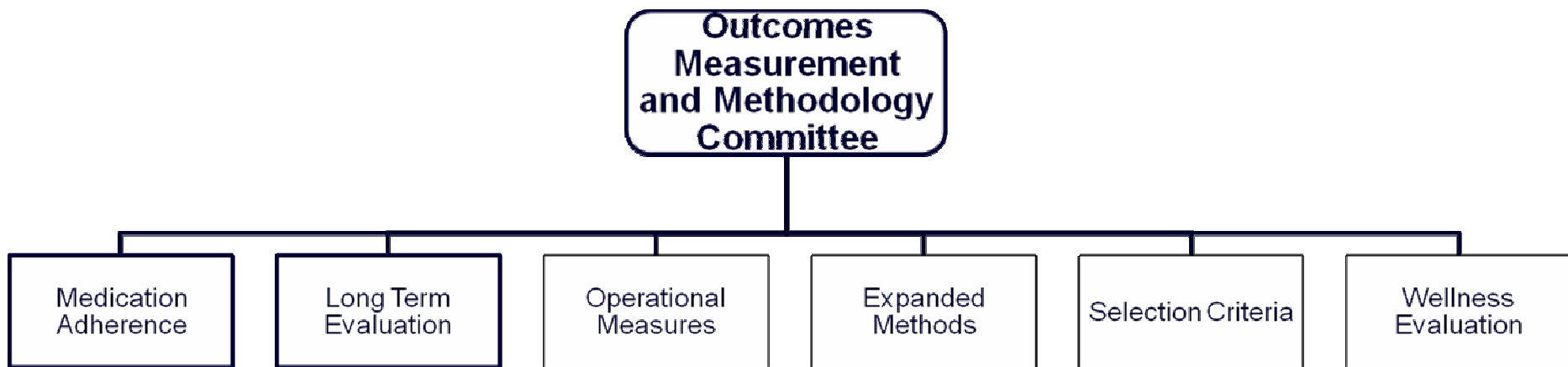
Outcomes Guidelines Project 2009

Outcomes Measurement & Methodology Committee

Volume IV Recommendations

dmaa







Recommendations

- Medication Adherence
 - Refinement of 2008 Measure of Medication Possession Ratio
 - A Measure of Persistence
- Long Term Evaluation
 - Key Concepts
- Operational Measures
 - Refinement of 2008 Operational Flow Diagram



Recommendations cont.

- Expanded Methods
 - Definition of Population Health Management
 - PHM Program Framework
 - Population Health Management Capabilities
 - Program Evaluation Core Measures



Recommendations cont.

- Wellness Evaluation
 - Use of Control/comparison group
 - Primary Outcome Measures
 - Timeline
 - Other Outcomes Measures
 - Use of Claims Data
- Selection Criteria
 - Testing Update and Preliminary Findings



Medication Adherence

Recommendation: Refining MPR

- The 2008 recommended measure for MPR is being refined as follows:
 - Numerator (Days Supply) \leq Denominator (Days)
 - Persistent Asthma Added as Condition
 - Long-term control drug classes per NAEPP guidelines: inhaled anti-inflammatory agents (cromolyn sodium, nedocromil), inhaled corticosteroids, oral leukotriene modifiers, inhaled long-acting beta agonists, and oral xanthines (methylxanthines)
 - A minimum of two claims, per member, for a specific drug class must be incurred to include the member in the calculation



Medication Adherence

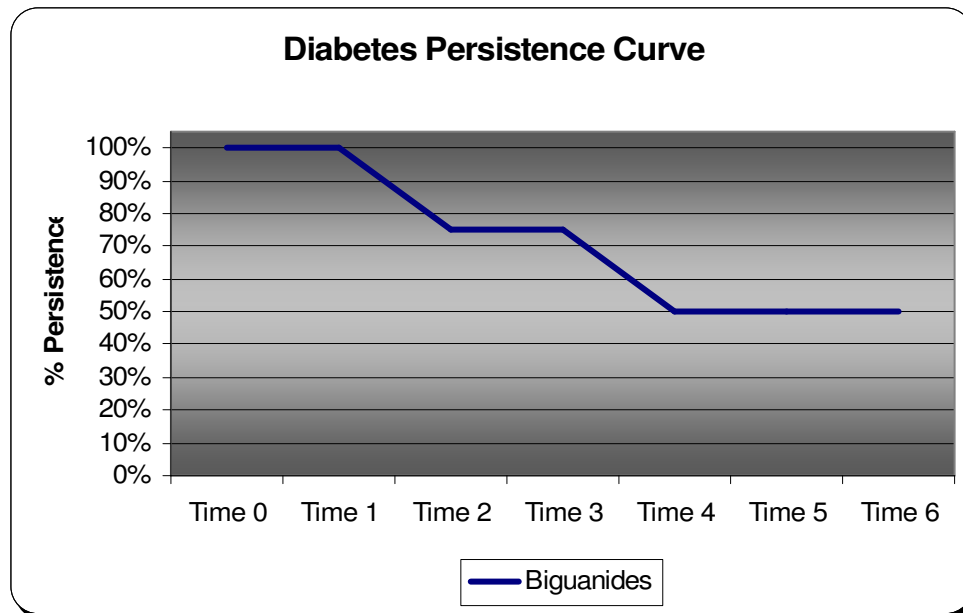
Recommendation: A Measure of Persistence

- Medication persistence can be defined as the “amount of time that an individual remains on chronic drug therapy”
- Inclusion/Exclusion Criteria Identical to MPR
 - **Except: “A minimum of two claims, per member, for a specific drug class must be incurred to include the member in the calculation” does not apply to persistence**
- Reported by Condition/Drug Class (as MPR)
- Reported as a Percentage Over Time Periods
- Visually Represented as a Kaplan Meier Curve
- 60-Day Permissible Refill Gap

Persistence Example

	Time 0	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6	
Mbr1	1	1	1	1	1	1	1	Eligible at year start; 100% persistent
Mbr2	1	1	1	1	0	0	0	Eligible at year start; 50% persistent
Mbr3	1	1	1	1	--	--	--	Eligible mid-year; 100% persistent
Mbr4	1	1	0	0	--	--	--	Eligible mid-year; 33% persistent
...								
P%	100%	100%	75%	75%	50%	50%	50%	

Time 0 = Initiation of Therapy, regardless of calendar date @ start
-- = "Null" entry (i.e. non-measurable data point due to start date)





Long Term Evaluation

Recommendation – Key Concepts

- (1) There are two different issues that affect measurement of impact over the long term
 - Baseline comparison: over time, measuring against a static baseline exacerbates any variation that might have affected the initial measured trend differences and masks any ongoing impact of non-program related factors on trend differences
 - Over time, correlated with the rate a program extracts any potential savings (which are finite) within a diseased population, the measured savings will decrease



Long Term Evaluation

Recommendation – Key Concepts

Baseline Comparison Issue

- (2) The issue of baseline comparison impacting long term evaluation comes into play after a population has been managed for over three years.
- (3) An alternative to measuring improvement over time against a static baseline could be to measure year over year impact or to re-baseline.



Long Term Evaluation

Recommendation – Key Concepts

Extraction of Potential Savings Issue

- (4) This concept applies to both populations managed by the same provider or populations that have been managed by one provider and then switched to a new program provider.
- (5) The issue of extraction of savings impacting long term evaluation is relevant to most programs and most measures used to evaluate these programs with the possible exception of complex case management.



Long Term Evaluation

Recommendation – Key Concepts

- (6) Because of the impact of the extraction of savings on their measurement over time, closely evaluating operational measures, clinical results, utilization management and intangible factors such as quality of life indicators are important additional components of evaluating program effectiveness.



Operational Measures

Recommendation: Refinement of the Diagram

- **Highlights** (please see word document for full diagram—
Operational_Measures_Refined_Flow_Diagram.doc)
 - Not Targeted box renamed - Not Targeted for Intervention
 - Targeted box split – Targeted for Interactive Coaching and Health Education and Targeted for Health Education Only
 - Enrolled Boxes renamed – Enrolled Through Opt-In Program and Enrolled Through Opt-Out Program
 - Participating box deleted

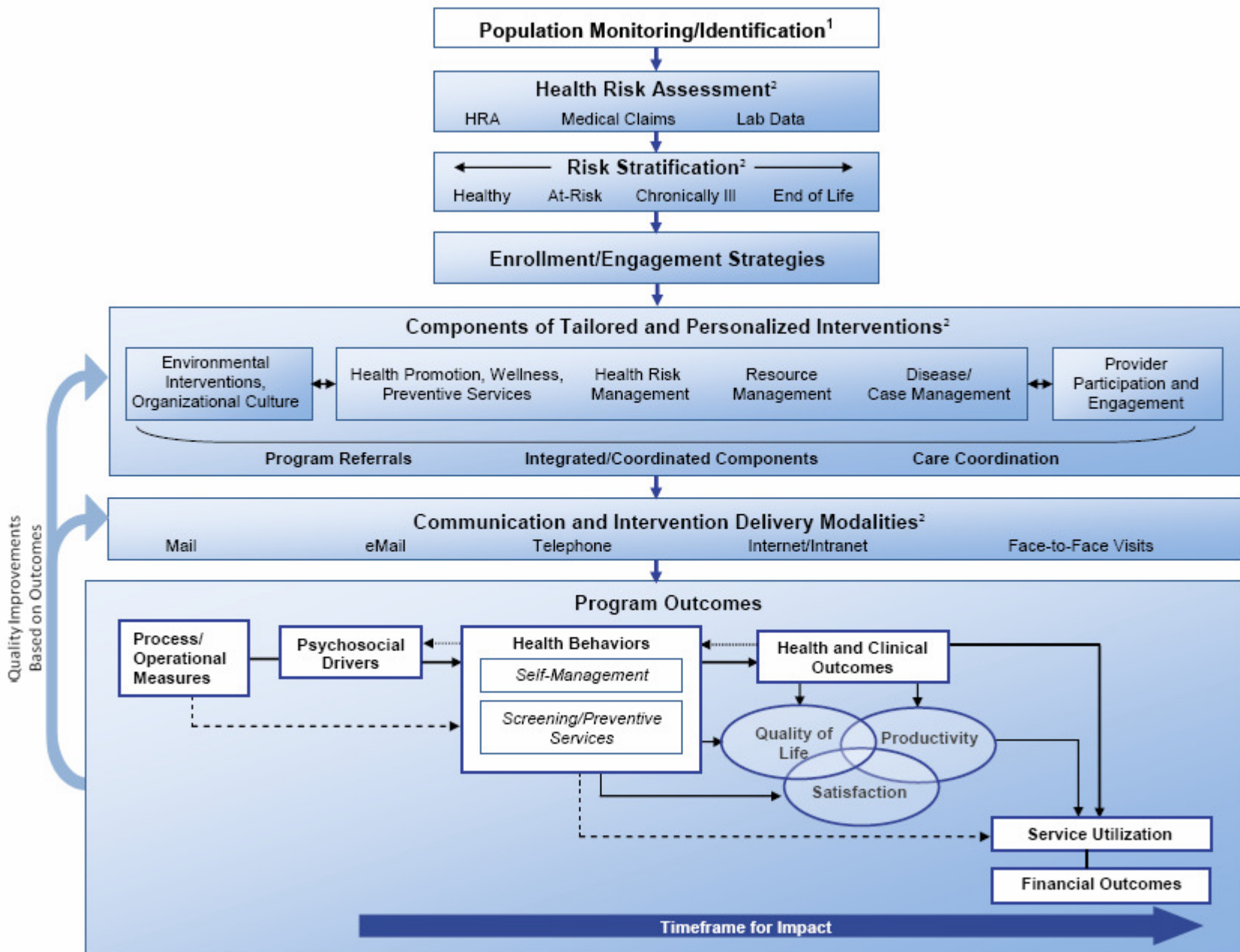


Recommendation:

Population Health Management Program Draft

Definition

- A Population Health Management program strives to address health needs at all points along the continuum of health and well being, through participation of, engagement with, and targeted interventions for the population. The goal of a population health management program is to maintain and/or improve the physical and psychosocial well being of individuals through cost effective and tailored health solutions.



¹For a more detailed discussion of monitoring and identification flow please refer to the work of the Operational Measures Workgroup

²Represents example components for each Essential Element. Does not necessarily reflect the universe of components.

PHM Capabilities

Capabilities	Basic Capabilities	Enhanced Capabilities
Engagement Strategies	Active or Passive	Combination of Active and Passive
Health Risk Assessment	Single data source	Multiple data sources
Risk Stratification	Single dimension	Multiple dimensions
Tailored Interventions	Population oriented	Member oriented
Program Components	Multiple components along the continuum that address needs specific to healthy, at risk, and/or chronically ill	Multiple Components Integrated and Coordinated
Program Delivery Modalities	Single Modality	Multiple Modalities
Information Exchange with Providers	One way - Outbound	Two way – Inbound and Outbound
Outcomes Measurement	Evaluate Effect on Outcome Domains (separately)	Assess Root Cause Impact on Outcomes (such as individual behavior changes)



Recommendation: PHM Program Evaluation Core Measures

- Cost and utilization
 - Medical costs
 - Health care utilization – Appropriate use
- Health and well-being
 - Health risks/behaviors
 - Quality of life
 - Health status
 - Productivity
 - Psychosocial Drivers
- Other
 - Program satisfaction
 - Process/operations measures



Wellness Evaluation

Recommendation: Use of Control or Comparison Group

- The use of an appropriate control/comparison group should be determined by the size and scope of the program being evaluated, by purchaser preferences and should reflect the understanding of the strengths and limitations of the various evaluation methodologies shown in Table 1.

Table 1. Control/Comparison Groups Methods for Wellness Program Evaluation

METHOD TO DEVELOP A COMPARISON GROUP	Randomized Control	Non-Randomized Control		Comparison Group		Own Control
		Matched Control	Un-Matched Control	Book-of-Business	Comparable Employee Population	
General description of method	Intervened population compared with individuals randomly selected to have services withheld	Intervened population compared with non-participants matched to have similar demographic, clinical and behavioral characteristics	Intervened population compared with non-participants	Intervened population compared with book-of-business population in same program	Intervened population compared with similar employer group(s) in same program	Pre/post comparison of intervened population.
Comparison time frame	Concurrent to intervention	Concurrent to intervention	Concurrent to intervention	Prior period	Prior period	Concurrent to intervention
Primary outcome measures	Modifiable risk factors, biometric variables	Modifiable risk factors, biometric variables	Modifiable risk factors, biometric variables	Modifiable risk factors, biometric variables	Modifiable risk factors, biometric variables	Modifiable risk factors, biometric variables
Population selection bias	None	Somewhat significant	Significant	None	None	
Source of comparison group	Population for whom program was implemented, randomly selected group withheld from program	Population for whom program was implemented, purchaser decision to not participate	Population for whom program was implemented, purchaser decision to not participate	Vendor data	Vendor data	Intervened population
Credibility of causal statements	Extremely strong	Moderate	Poor	Very poor	Very poor	Very poor
Control of confounding variables	Controls known and unknown, measured and unmeasured confounding variables.	Controls known and measured confounding variables.	Does not control confounding variables	Does not control confounding variables	Does not control confounding variables	Does not control confounding variables
Program sponsor resistance to approach	High	Moderate/High	Moderate	None	None	None
Ease of implementation	Very difficult	Difficult	Somewhat Difficult	Easy	Easy	Easy
Clarity of method to lay audience	Very clear	Very unclear	Somewhat unclear	Clear	Clear	Clear

Table 1 (cont'd)

Multiyear application vs. single year application	Much harder	Much harder	Somewhat harder	Same	Same	Same
Method availability	Rarely possible	Occasionally possible	Occasionally possible	Always possible	Usually possible	Always possible
Bleed of interventions to comparison population	Control group may get provider-based interventions, other vendor interventions, secular interventions and self-care	Control group may get provider-based interventions, other vendor interventions, secular interventions and self-care	Control group may get provider-based interventions, other vendor interventions, secular interventions and self-care	Comparison population already intervened	Comparison population already intervened	NA
Key strengths	Gold standard evaluation method	Possible to infer reasonable level of causality without experimental design	Least costly and easiest method of compared intervened to non-intervened population	Low cost, availability of data, ability to answer question asked: (How does my group compare to other groups that have been in the program?)	Low cost, availability of data, ability to answer question asked: (How does my group compare to other groups that have been in the program?)	Low cost, availability of data, ability to answer the question asked: (Did the intervened group change over time.)
Key problems/biases	Cost, sponsor resistance, IRB imperatives, low generalizability	Cost, availability of control group HRA data	Limited ability to make causal inferences,	No inferences on causality are possible	No inferences on causality are possible	No inferences on causality are possible; No information on relative performance of intervened group.



Wellness Evaluation

Recommendation: Primary Outcome Measures

- The primary outcome measures of wellness program evaluation should be modifiable behavioral risk factors and related biometric variables. The minimum set of these outcome measures may include:
 - Diet
 - Exercise/physical activity
 - Tobacco use
 - Alcohol/drug use
 - BMI
 - BP
 - Blood glucose
 - Stress levels
 - Lipid profile



Wellness Evaluation

Recommendation: Timeline

- Measurement of primary outcome measures should be made at baseline.
- Re-measurement of primary outcome measures should take place at 12-month intervals.



Wellness Evaluation

Recommendation: Other Outcome Measures

- Quality of life should be measured at baseline. Such measures can be made via single questions or by validated QOL instruments. [Examples of each to be provided in full guideline.]
- Both presenteeism and absenteeism should be measured at baseline. [Discussion and examples of instruments to be provided in full guideline.]
- Re-measurement of these outcomes measures should be made at 12-month intervals.



Wellness Evaluation

Recommendation: Use of Claims Data in Program

Evaluation

- DMAA recommends that when claims data are to be used to establish either a financial baseline or a utilization baseline the following criteria should be met:
 - Program scope and cost are commensurate with such an analysis
 - The claims data are routinely available
 - There is a minimum of 3 years of program exposure
 - The sample size is adequate
 - There is a means of attributing affects
 - Appropriate claims analysis methodology is established



Selection Criteria: Testing Overview

- Purpose: test selection criteria for denominators for outcome metrics (e.g., clinical, financial) to be used for program comparisons.
- Test selection criteria developed in 2008 (published in Outcomes Guidelines Report Volume III pp 25-29) for:
 - Diabetes
 - Persistent asthma (adult and pediatric)
 - COPD
 - Heart Failure
 - CAD
- Determine and test appropriate eligibility and identification timeframe for meeting:
 - Accuracy and balance (are diseases identified in Time1 also found in Time2)
 - Fairness and relevance (overlap between DMAA and vendor criteria) ²⁵



Three Testers

- Number of Members Included in Study
 - 1.5 to 1.9 Million Members
 - Enrollment and claims over 2 1/4 to 4 1/4 years.
- Number of Members reduced due to:
 - Age & Continuous Eligibility Requirements
 - .5 to .7 Million Members included in measures.
- Testers Expected to Test:
 - Selection Criteria Accuracy and balance
 - Overlap between the DMAA and DM criteria.



Selection Criteria

Findings: Overlap of DMAA and Proprietary Criteria

- In general, DMAA criteria finds fewer people than proprietary criteria
- Overlap varies by condition
 - Diabetes & CAD – DMAA found 3% & 8% people that proprietary did not
 - CHF, Asthma and COPD – DMAA found 25 to 27% people that proprietary did not
- Overlap varies by proprietary method significantly
- Proprietary finds many people that DMAA did not.
- These results changed only slightly with added drug claim criteria.



Selection Criteria

Recommendation: Continue Testing

□ Recommendations

- Modify criteria for continued testing as follows for use as selection criteria in comparative clinical and financial performance analyses
- All 5 conditions
 - **Eligibility time period = 6 months**
 - **ID timeframe = 24 months with at least 3 months of run out (with standard run out defined for each type of analysis)**
- CAD
 - Additional qualifying criteria: one visit/op/ED encounter + **at least one claim for nitrate or lipid lowering drug**
- COPD
 - Additional qualifying criteria: one visit/op/ED encounter + **at least one claim for triotropium**



Selection Criteria

2010 Goal: Field Test Criteria

How well suited are the denominators for:

- Clinical measures.
- Financial measures.

Process:

- Find 6 to 8 Testers to run various tests.
- Gain NCQA insights in how to Field test.



Outcomes Agenda 2010





Outcomes and Methodology

- Selection Criteria
 - Complete testing and recommend modifications to the Outcomes Steering for 2010 Report
- Population Health Management
 - Continue work of 2009 to include evaluation methodology development
 - Expand through grant to include model analysis and possible demonstration
 - Work with HIT committee to develop health information technology overlay for the PHM framework
- Operational Measures
 - Complete measures and specifications for the 2010 report



Outcomes and Methodology cont.

- Self Report Medication Adherence and Barrier Identification
- Outcomes Guidelines Medicaid Application
- Data Repository
 - Develop partner relationship
 - Secure external funding
 - Secure 10 data contributors
 - Launch repository Late 2010