

# Guidelines for Evaluating Population Health Improvement Program Outcomes

---

*November 2008*



# Guideline Sections

---

- ❑ Methodological Considerations
- ❑ Measures/Metrics
- ❑ Wellness
- ❑ Methodology for Expanded Populations

# Methodological Considerations

# Evaluation Design

---

- Evaluation design: DMAA recommends the use of a pre-post study design with an internal or external comparison group that is equivalent to and assessed over the same time period as the group receiving the intervention. DMAA recognizes, however that a comparison group that is both equivalent and concurrent may not always be available in applied settings. Accordingly, DMAA recommends that evaluations using a pre-post study design without a comparison group make explicit efforts to control potential biases and error introduced by the design and that the potential impact of the design on the interpretation of the findings be made clear.



# Study Design Benefit Comparison

---

- Recognizing the need for clarification on study design methods, DMAA, as a follow-up to the study design recommendation developed in the Volume I Guidelines Report, has developed a study design matrix and recommends a review of the matrix information to gain a better understanding of study design methods/implications.
- Please refer to the Study Design Benefit Matrix.

# Population Identification Methodology

---



- DMAA recommends the Annual Qualification methodology as the method of choice for population identification

# Defining the Population



- Measurement period: Recommend one year for baseline and subsequent years
- Criteria for inclusion in measurement: Recommend that the commercial and Medicare member population be enrolled with buyer for  $\geq 6$  member months, Medicaid TANF  $\geq 1$  month
- Look back period: Recommend 12 months of measurement period as well as at least 12 months of the preceding period for the purpose of program evaluation

# Defining the Population *(cont.)*



- Defining a member month: DMAA recommends that a member month be defined as members enrolled on the 15th of the month for commercial and Medicare populations when possible.
- Claims Runout Period: DMAA recommends 3 months with completion or 6 months with no completion contingent upon consistent payment patterns.

# Exclusions/Outliers

---

- Exclusions: DMAA recommends that there should be three types of exclusions from the evaluation for financial and utilization measures:
  - Patients with conditions such as:
    - ESRD
    - HIV/AIDS
    - Transplants
    - Non-skin cancers with evidence in claims of active treatment
    - Hemophilia
  - Claims for diagnoses such as: (but not the person with these claims)
    - Trauma with hospitalization
    - Skin cancers
  - Stop-loss at member level; such as removing claims above \$100K annually, indexed to grow at future years concurrent with an appropriate trend



# Comparing Methods that Define Outliers

---

- Building on earlier work, DMAA has developed a matrix that reviews methods commonly used for defining outliers and recommends review of the matrix when designing DM evaluations.
- Please refer to the Methods to Define Outliers Matrix.

# Guiding Principles for Selecting Populations for the Purpose of Program Evaluation



- Selection criteria applies to both program evaluations and comparative performance evaluations;
- Selection criteria is consistent with recommendations already approved by the OSC;
- Selection criteria applies to all non-clinical measures. In addition, selection criteria for clinical measures will be synonymous with the specific denominator criteria of the endorsed clinical measures;
- The development of algorithms and the codes to implement them is an iterative process;

# Selection Criteria Overview



- Selection criteria are defined for:
  - Diabetes
  - Heart failure
  - Coronary artery disease (CAD)
  - Persistent asthma
  - Chronic obstructive pulmonary disorder (COPD)
- These criteria are to be used for comparative performance metrics. They are not intended to define participants for specific program interventions.
- The criteria are condition-specific. In 2009, the work group will consider approaches to identifying individuals with multiple conditions.

# Selection Criteria General Comments



- The condition-specific definitions are recommended for detailed testing in 2009. Testing may result in changes/additions.
- When possible these criteria match or are similar to those being developed and tested by NCQA for clinical measurement specifications for DM Programs
  - DMAA's criteria reflect the need for a "generic" denominator that can be used and/or tweaked for multiple purposes
  - Both DMAA and NCQA will be testing these specifications in the coming year with both organizations benefiting from the investigations
- Note that detailed specification of all codes except NDC codes are included in the Outcomes Guidelines Report Volume 3. Specific NDC codes for use with pharmacy claims will be provided on the NCQA website.

# Selection Criteria: Building Blocks



- Required benefits
  - All conditions require medical benefits
  - Diabetes and asthma also require pharmacy benefits for denominator definition
  - All conditions may require pharmacy benefits for some metrics' numerators
- Types of data
  - All conditions require access to professional claims or encounters and to facility claims. It must be possible to determine site or place of service.
  - Codes
    - CPT-4 and ICD-9 codes for professional claims/encounters
    - ICD -9, HPCPS, and revenue codes for facility claims (outpatient/inpatient)
    - NDC codes for pharmacy claims

# Selection Criteria: Building Blocks *(cont.)*



- Minimum eligibility (covered or insured status)
  - Same definition across all conditions
  - Whether this requirement may differ for clinical versus financial metrics remains an open question.
  
- Demographics
  - Age as of last day of measurement period
  - Upper limit
    - Diabetes, heart failure, CAD and COPD have no upper age limit
    - Upper age limit for asthma is 56.
  - Lower limit
    - 18 for diabetes, heart failure and CAD
    - 5 for asthma
    - 40 for COPD

# Selection Criteria: Building Blocks *(cont.)*



- Identification Time Frame
  - The period during which all identification criteria must be met to qualify for selection
  - The criterion is under consideration for next year.
- Type and number of codes that must occur in identification time frame
  - See detailed specifications in Outcomes Guideline Report Volume 3
  - Example for heart failure
    - 1 or more acute inpatient discharge with listed ICD9 codes in any position OR
    - 2 or more office/outpatient visits or encounters or emergency department visits at least 14 days apart with listed ICD9 codes in any positions for listed CPT codes for eligible visits

# Selection Criteria: Building Blocks *(cont.)*



## □ Exclusions

### ■ Condition-specific exclusions

- Diabetes: gestational diabetes
- Asthma: COPD

### ■ Condition-independent exclusions

- Next year we will consider common exclusions, including codes indicating residential treatment, such as hospice and other exclusions noted in DMAA Outcomes Guidelines Report Volume 1

# Trend



- DMAA recommends the use of a non-chronic population for the purpose of calculating trend.
  - For this purpose, the non-chronic population is defined as those members not identified as having any of the following “common chronic” conditions: diabetes, CAD, heart failure, Asthma, COPD.
  - In addition, when warranted and mutually agreed by the parties, members with certain other conditions may be excluded from the non-chronic population if these conditions are also being managed by another DM program outside of the five “common chronics” listed previously.

# Calculating Trend *(cont.)*

---



- DMAA recommends the use of the average difference between historical chronic and non-chronic trends to adjust current year non-chronic trend.

# Risk Adjustment

---



- Risk adjustment: DMAA recommends that parties must agree on mutually acceptable risk adjustment method, ideally a commercially available tool

# Risk Adjustment (*cont.*)



- In deciding (1) whether and (2) how to approach risk adjustment for a particular DM program for a specific population, it is useful to categorize outcomes of interest into one of the following 2 categories:
  - Category 1: Those believed to be impacted only by exogenous confounders and not the DM interventions, where there is no concern that program impacts will be altered by risk adjustment (e.g., non-chronic trend)
    - For this category of variables, one should utilize an appropriate risk adjustment method, ideally a commercially available risk adjustment tool or other non-proprietary validated method.

# Risk Adjustment (*cont.*)

---

- Category 2: Those believed to be impacted by exogenous confounders as well as by program interventions that potentially may be inappropriately distorted or discounted by risk adjustment (e.g., condition prevalence or severity, casemix)
  - For this category of variables, the next step is to examine the potential magnitude and importance of the potential exogenous confounder(s)
  - If the potential magnitude is large and/or highly important, then one must consider which available risk adjustment methods permit a reasonable job of adjusting for the offending confounders without seriously distorting or discounting program impacts

# Small Populations/Sample Sizes

---

- If more than one method is available, then the one with the least likelihood of distorting program impacts while reasonably adjusting for confounding factors is preferred
- In some cases, using a “minimalist” approach, such as age-gender or simple prevalence adjustment, may be more suitable than more complex risk adjustment tools
  - This is because the more comprehensive or explanatory the risk adjustment method, the greater the likelihood that some of the input variables for that method are factors that are positively impacted by DM programs.

# Small Populations/Sample Sizes (*cont.*)



- Dealing with small sample sizes: DMAA advises that as population size drops below a certain level, calculated DM financial outcomes begin to lose credibility and reliability. (This level exists on a continuum, is empirically determined, and can be estimated using common actuarial practices or statistical power methods.)
  - The smaller the population falls below this level, the more that random variation will influence results and interfere with the credibility and reliability of the calculated outcome.
  - As mutually agreed, parties may prefer to avoid this concern by choosing not to calculate financial outcomes for such small populations, or may elect to mitigate this concern by using a credibility factor approach to blending their smaller population result with some larger (typically, comparable “book of business”) population to increase the credibility of this result.

# Small Populations/Sample Sizes (*cont.*)



- Recommended use of three methods for dealing with small populations
  - Blend results (using standard medical cost savings methods) for the small population with results from “book of business” or larger reference population.
  - Using a “book of business” or larger reference population results, derive a factor that estimates the percentage of total medical costs saved per member who receive a significant level of support. Use this factor to calculate savings for all members in small population receiving this level of support.
  - Use “book of business” or a larger reference population’s results to build a statistical model that assesses all of the factors that drive savings. Apply the model to the smaller group’s data to calculate savings estimates.

# Small Populations/Sample Sizes (*cont.*)

## Alternative 1



**Blend results (using standard medical cost savings methods) for the small population with results from “book of business” or larger reference population.**

- Blends customer specific results with results from a more stable “book of business” or larger reference population that is assumed to be comparable, typically without severity, age, sex or similar adjustments
- Is in line with standard actuarial processes for rating, which blend the computed premium calculated by book of business and the premium calculated by small group experience rating in a ratio that more heavily weights the book of business at smaller sample sizes.
- Enables the results of a small population to be factored in to the ultimate result giving some representation to the contribution of the small population.

# Small Populations/Sample Sizes (*cont.*)

## Alternative II



**Using a “book of business” or larger reference population results, derive a factor that estimates the percentage of total medical costs saved per member who receive a significant level of support. Use this factor to calculate savings for all members in small population receiving this level of support.**

- Enables group level of activity and perhaps cost data to be utilized in deriving savings.
- Uses a more stable “book of business” or larger reference population results.

# Small Populations/Sample Sizes (*cont.*)

## Alternative III



**Use “book of business” or a larger reference population’s results to build a statistical model that assesses all of the factors that drive savings. Apply the model to the smaller group’s data to calculate savings estimates.**

- Enables group level information to be utilized to derive savings.
- Uses more stable “book of business” or larger reference population results.
- May build upon other studies such as experiments conducted in the medical literature

# Small Populations/Sample Sizes (cont.)



- The following example demonstrates the significance of the impact of population variability on medical cost savings assessments.
- The table shows confidence intervals on first year per member per month (PMPM) medical costs savings estimates in a population with a robust chronic condition management program in place.
- The savings estimates were derived from methods compliant with DMAA's Outcomes Measurement Guidelines.
- The variance was estimated using repeated samples from a large commercial health plan population in which a chronic condition management program was implemented.

Population with Chronic Conditions*	Total Population	Total Population PMPM Medical Cost Savings** (95% Confidence Intervals)
30	500	-\$17.29 to \$29.20
60	1,000	-\$21.79 to \$35.50
120	2,000	-\$11.07 to \$24.02
180	3,000	-\$7.02 to \$17.77
240	4,000	-\$3.91 to \$15.24
300	5,000	-\$3.74 to \$14.71
600	10,000	-\$2.88 to \$13.26
900	15,000	\$0.24 to \$10.24
1,200	20,000	\$0.74 to \$10.00
1,500	25,000	\$1.43 to \$9.04
1,800	30,000	\$1.98 to \$8.86
2,400	40,000	\$2.06 to \$8.36
3,000	50,000	\$2.53 to \$8.09

\*\*Savings for the population with chronic conditions is divided by the total population member months

# Small Populations/Sample Sizes (cont.)



- For example, in the first line, 500 samples of members were pulled from the entire population generating, in each case, 30 members with chronic conditions.
- The medical cost savings algorithm was applied to each of these samples and the variance in these results was measured to produce the confidence interval cited below. This process was repeated for different-sized samples up to 3,000 members with chronic conditions.

Population with Chronic Conditions*	Total Population	Total Population PMPM Medical Cost Savings** (95% Confidence Intervals)
30	500	-\$17.29 to \$29.20
60	1,000	-\$21.79 to \$35.50
120	2,000	-\$11.07 to \$24.02
180	3,000	-\$7.02 to \$17.77
240	4,000	-\$3.91 to \$15.24
300	5,000	-\$3.74 to \$14.71
600	10,000	-\$2.88 to \$13.26
900	15,000	\$0.24 to \$10.24
1,200	20,000	\$0.74 to \$10.00
1,500	25,000	\$1.43 to \$9.04
1,800	30,000	\$1.98 to \$8.86
2,400	40,000	\$2.06 to \$8.36
3,000	50,000	\$2.53 to \$8.09

\*\*Savings for the population with chronic conditions is divided by the total population member months

## Small Populations/Sample Sizes (*cont.*)

---

- Measuring Savings in Small Populations
  - Measuring medical cost savings in small populations poses significant challenges to the disease management industry. In small populations, the effects of even a very small number of large claimants can make interpretation of medical costs savings assessment results difficult. In general, medical cost data are highly variable; in large populations, the variability in the data becomes much easier to manage than it is in small populations.
  - Caution is advised in producing medical cost savings measures using these methods in populations with small numbers of members receiving management for chronic conditions. High variability will frequently result in conflicting, misleading or inaccurate indications of program-related impact.

# Measure Sets

# Financial Metrics



- Financial metric: DMAA recommends using health care cost outcomes as the metric for assessing the financial impact of the program.
  - Health care cost outcomes would be measured using medical and pharmacy claims (where available) to calculate changes in total dollars, and also expressed as per-member-per-month charges.
  
- Which costs to use: Recommend using paid and/or allowed costs

# Utilization



- DMAA recommends the following measures as a minimum set of useful utilization metrics to complement and corroborate the financial outcomes measures for DM program evaluation:
  - All-cause admission rates per thousand members per year and ER visit rates per thousand members per year for the “diseased” or eligible population;
  - Condition-specific admission per thousand members per year and ER visit rates per thousand members per year for the entire insured or covered population (using principal diagnosis only to define condition-specific events).

# Clinical Metrics



- 
- Clinical Metrics: DMAA has incorporated, as a starting point, the series of ICD-9 codes established in the latest version of the DMAA *Dictionary of Disease Management Terminology*.

In 2007, DMAA will develop standardized identification criteria for defining both the numerator and the denominator for an agreed list of disease-specific clinical metrics for the purpose of program evaluation.

# **Guiding Principles for Evaluation and Comparability for Clinical Effectiveness Measures**

# Comparative Performance Reporting

---



- DMAA recommends that comparative performance reporting should enable comparisons among organizations offering disease management programs in addition to standard program evaluations.

# Performance Results Reporting



- DMAA recommends that book of business performance results be reported with analysis appropriate to context and relevance, e.g., at the geographic, industry or market level, as is feasible.

# Data Sources



- DMAA recommends that organizations wishing to participate in comparative performance reporting secure consistent access to relevant transactional data and patient self-reported data to compute performance results;
- DMAA recommends that patient self-reported data be collected under highly standardized and scientifically rigorous circumstances to improve reliability and validity (e.g., standardized patient survey).

# Eligibility



- DMAA recommends that members become denominator-eligible if they have been identified as eligible for the intervention (intent-to-treat principle);
- We recommend that members eligible for denominator have had the opportunity to be exposed to the DM program for no less than three months;
- If required data elements are not available for all otherwise eligible members, it must be demonstrated that no material bias exist for the calculation of performance results (see Section on Validation);
- To the extent that sample-based measurement is performed, DMAA recommends that a representative sample be drawn from the sampling frame (e.g., all denominator eligible members);
- DMAA recommends that required sample sizes are calculated to allow performance estimates to range  $\pm 10\%$  within a 90% confidence interval.

# Risk Adjustment



- DMAA recommends that no statistical adjustments are performed for such variables in order not to mask important potential disparities that may exist in the provision of health care services and results;
- DMAA recommends that relevant reporting strata are created to understand that potential differential impact of DM programs on members of different age, gender, ethnicities, etc. Stratification by insurance product members enjoy can also be helpful (commercial, Medicaid, Medicare);
- DMAA recommends that when feasible (e.g., acuity/severity can be reliably/validly assessed and is easily available), performance results be either stratified or adjusted for pre-intervention differences.

# Sample Size



- DMAA recommends that performance results are not publicly reported when the population size is relatively small (and parametric statistical techniques are not applicable or begin to break down).
- DMAA recommends that sample-based and publicly reported performance results (e.g., sampling from the population to solicit patient self-reported data) be generally based on no less than N=100 responses (change previous slide to 90 percent +/- 10 percent).

# Calculating Results



- DMAA is interested in reporting results that show absolute level of performance and improvement in performance over time.

# Verification

---



- DMAA recommends that publicly reported, comparative performance results pertaining to DM programs or organizations are verified.

# Group I Clinical Metrics for Program and Comparative Performance Measurement

# Group I Measures Recommendations



- These are examples of parameters for clinical evaluation that may not be relevant to all and for which data may not be available.
- When medication prescriptions are referred to, it should be in the context of relevance and appropriateness to the patient.
- These are subject to update.

# Group I Measure Categories

## Asthma

- Flu vaccination
- Pneumococcal vaccination
- Smoking cessation ID & Advice
- Current medication use
  - Controller medication

## COPD

- Flu vaccination
- Pneumococcal vaccination
- Smoking cessation ID & Advice
- Spirometry evaluation
- Medication use
  - Bronchodilator

## Heart Failure

- Flu vaccination
- Pneumococcal vaccination
- Smoking cessation ID & Advice
- Medication (Persistence)
  - BB
  - ACE/ARB
  - Anticoagulants (with chronic or proxysmal AF)\

## Coronary Artery Disease

- Flu vaccination
- Pneumococcal vaccination
- Smoking cessation ID & Advice
- LDL testing & control
  - annual test
  - LDL < 100; > 130
- BP
  - BP <= 140/90
- Medication(Persistence)
  - BB
  - ACE/ARB
  - Aspirin

## Diabetes

- Flu vaccination
- Pneumococcal vaccination
- Smoking cessation ID & Advice
- Daily aspirin use
- LDL testing & control
  - annual test
  - LDL < 100; > 130
- HbA1c testing & control
  - annual test
  - HbA1c < 7.0; > 9.0
- BP
  - BP <= 130/80; > = 140/90
- Eye Exam
- Nephropathy testing

## Self Management Definition

---

- Self-management consists of the on-going processes and actions taken to manage/control one's own condition, with the goal of improving clinical outcomes, health status and quality of life.

# Medication Adherence

---



- Identified in Outcomes Guidelines Report Volume 2 as a possible measure of clinical performance.
- In 2008, developed a detailed specification for the medication possession ratio as one measure of medication adherence.

# Medication Adherence Recommendation for Measure of MPR

---



- Medication Possession Ratio (MPR) – Operationalized definition of adherence – retrospective assessment of the number of doses dispensed in relation to the dispensing period

Reference: Cramer JA, Roy A, Burell A, et al. Medication Compliance and Persistence: Terminology and Definitions. *Value Health* 2008 2(1): 44-47

# Medication Adherence

## Recommendation for MPR Specifications



- MPR is a population-based measure reported as a percentage
- Data sources for the measure are administrative pharmacy claims and eligibility data
- Evaluation period is defined as a fixed calendar length – 12 months (annual)
  - 1 month claims run-out to capture lag
- Enrollment criteria: a continuous evaluation period with no more than a 45 day gap in pharmacy benefits coverage
- Denominator is the duration from first (index) prescription to the end of the evaluation period
- Numerator is the days supplied over the same period

# Medication Adherence Recommendation for MPR Specifications -- Continued



- Reporting MPR will include:
  - Quartiles
- MPR will be reported by condition, by drug classes applicable to that condition
- Individuals with multiple conditions (e.g. CAD and Diabetes) will be counted for all conditions, for all appropriate drug classes

# Medication Adherence

## Recommendation for MPR Inclusions/Exclusions



- ❑ Measure is intended for oral medications only
- ❑ Measure is intended for more prevalent common chronic conditions (CAD, CHF, Diabetes, Hypertension, Hyperlipidemia)
- ❑ Measure excludes inhalers and liquids
- ❑ Methodology excludes “carry in” from prior evaluation period
- ❑ Methodology excludes “carry out” – when medication supply goes beyond evaluation period
- ❑ Concomitant therapy, contraindications, medication switches, samples, and overlap are not covered by the methodology. These, as well as utilization confounders, will be highlighted in methodology
- ❑ To be included, index prescription must occur within the first 6 months of the evaluation period

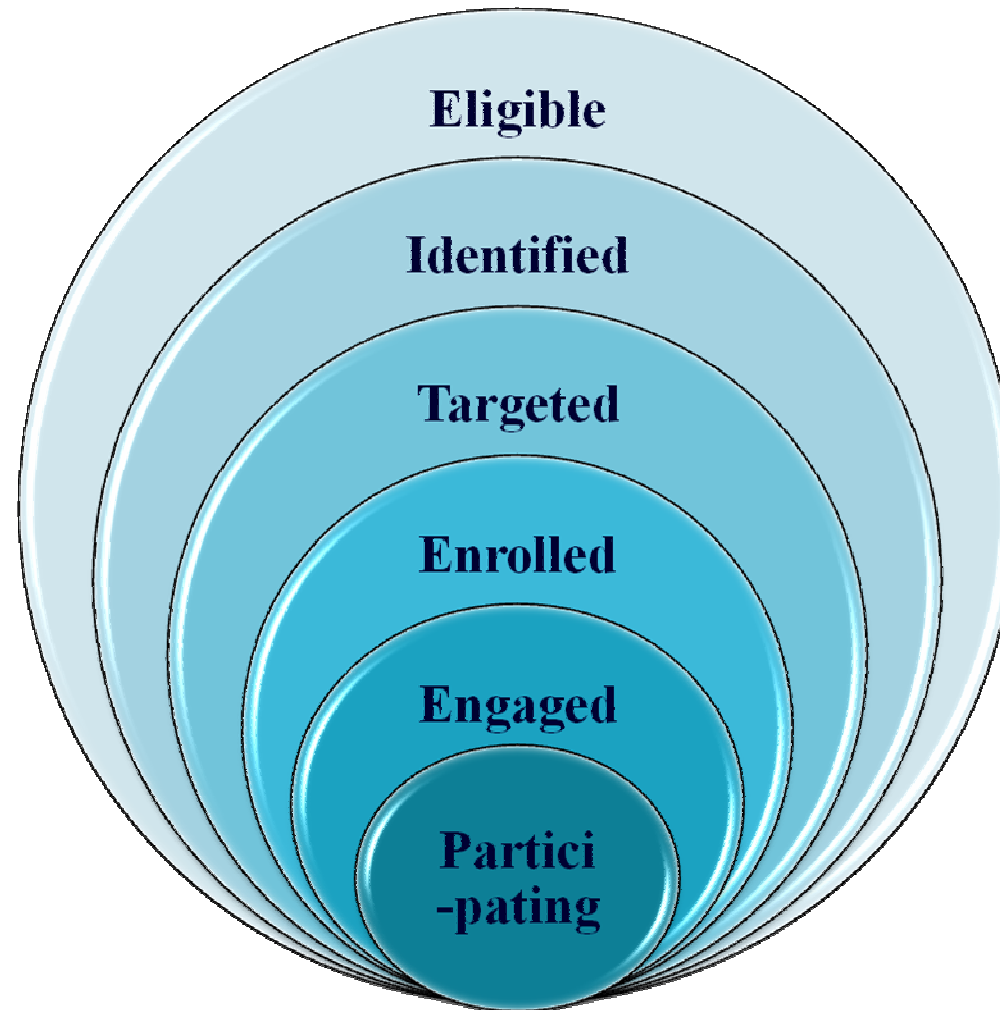
# Operational Processes

---



- Operational process flow
- Process stage definitions
- See Outcomes Guidelines Report Volume III for possible measures for each stage of processes

# Operational Measures Recommendations for Process Flow Diagram



# Operational Stages: Recommended Definitions



- Total Program Eligible Population  
Individuals meeting benefit eligibility and eligibility requirements established by the plan to be considered for the DSM care management program
  - Identified Population  
DSM program-eligible members who are identified as having a qualified condition
  - Targeted Population  
Members who are identified as having a qualified condition, and who are targeted for program intervention.

# Operational Measures: Recommended Definitions



- **Enrolled Population**

*For Opt-in program:* members consenting to participate in the program

*For Opt-out program:* targeted population considered enrolled, unless action taken to disenroll

- **Engaged Population**

*Passive engagement* strategies do not require active member consent, or acknowledgement of program participation

*Active engagement* strategies do require active member consent, or acknowledgement of program participation

- **Participating Population**

*For opt-in program:* Enrolled members who are actively participating in DSM program, and typically broken down by stratification level and/or level of engagement and intervention.

*For opt-out program:* Enrolled members who have not taken action to discontinue intervention, and typically broken down by<sup>57</sup> stratification level and/or level of engagement and intervention.

# Additional Metrics



- Consider use of one of the SF tools (e.g., SF-8™, SF-12®), SF-36®) to measure general mental and physical health status.
- Consider assessing participant satisfaction using the DMAA Standardized Participant Satisfaction survey



---

# Wellness

# Wellness Evaluation Methodology

---



- Goal is to develop recommendations for evaluation strategies for wellness programs that are:
  - Consistent with DM recommendations AND
  - Appropriate for wellness programs
- Focused this year on comparing DM and Wellness programs on key factors

# Wellness Program Definition



- Wellness programs are designed to help individuals maintain and improve their current levels of health and well-being by identifying health risks and assisting individuals to both understand and mitigate these risks. Wellness programs are designed to increase awareness of factors that can affect health and longevity and to enable individuals to take increased responsibility for their health behaviors. Wellness programs are designed to prevent or delay the onset of disease as well as to promote healthy lifestyles and general well-being. Effective wellness programs employ a variety of behavior change techniques and lifestyle management strategies.

# Wellness Program Definition (*cont.*)

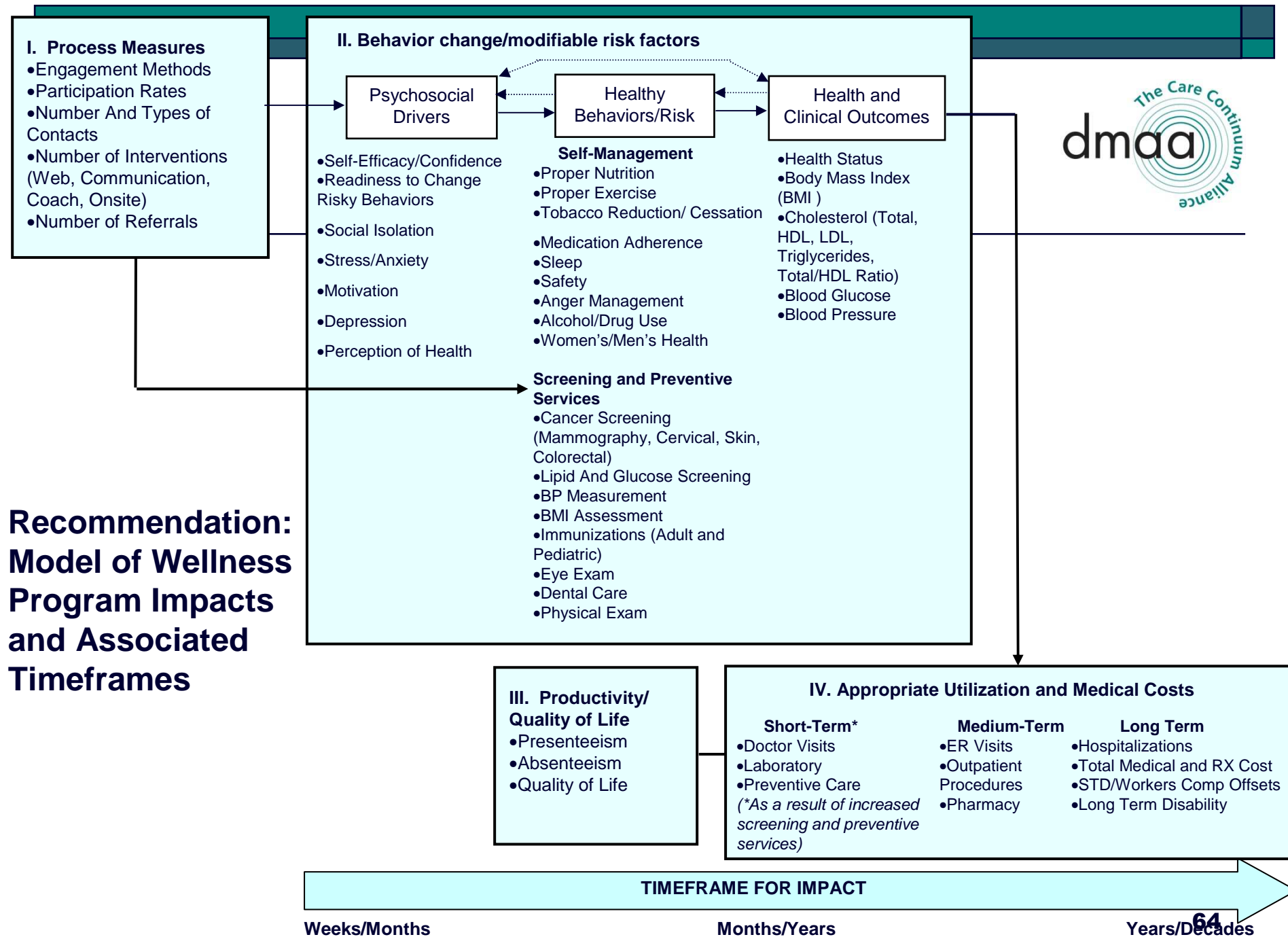


- The following are examples of components of wellness programs (note that this list is not exhaustive):
  - HRA
  - Biometric screening (BP, cholesterol, etc)
  - Smoking cessation
  - Weight loss
  - Diet & nutrition
  - Stress reduction
  - Exercise, fitness programs
  - Ergonomic programs
  - Safety (workplace and home)
  - Sleep hygiene
  - Health advocacy
  - Disease screening
  - Immunization

# Wellness Program Definition (*cont.*)



- Wellness programs target the total population and participation is not primarily driven by disease state. This approach differs from a total population disease management approach that could offer programs across the full health spectrum including both wellness and disease-specific components.



**Recommendation:  
Model of Wellness  
Program Impacts  
and Associated  
Timeframes**

# Comparison of Current State of DM /Wellness Programs



	<b>DM Program</b>	<b>Wellness Program</b>
<b>Health status of target population</b>	Always have chronic illness	Total population which varies from optimal health to chronically ill
<b>Primary Focus of Intervention</b>	Optimization of chronic condition management	Modification of health risk behaviors
<b>Primary Outcome Metrics of Interest</b>	Health care costs and utilization rates (hospitalizations, ER visits and use of specific diagnostic and therapeutic procedures)	Specific health risk behaviors , indirect health related costs (productivity, etc)
<b>Unit cost of program</b>	Moderate	Low to moderate
<b>Claims data availability (for purposes of evaluating program impacts)</b>	Usually available	Variable availability, less likely than with DM program
<b>Time frame for impact on participant behavior</b>	Near term	Near term
<b>Time frame for impact on health status</b>	Near to medium term	Near to long term depending upon program elements and aims

# Comparison of Current State of DM /Wellness Programs



Time frame for impact on health care costs	Near to medium term	Medium to long term
Use of health risk assessment for baseline assessment	Variable, less likely than with wellness program.	Nearly always
Use of biometric screening data for baseline assessment	Variable	Variable
Status/availability of narrative coaching/encounter notes for analysis	Variable	Usual (depending upon program type)
Availability of process measure metrics	Always	Always
Approach to evaluating financial impact of program	Direct analysis of claims data	Modeling of risk factor changes and their effects on health care costs, productivity and direct analysis of claims data
Availability/use of case reports	Usually available and appropriate to <i>illustrate</i> program effects, if not to <i>measure</i> them	Usually available and appropriate to <i>illustrate</i> program effects, if not to <i>measure</i> them

---

# Methodology Guidelines for Evaluating Expanded Population Programs

# Expanded Methodology

## General Guiding Principles

---



1. Understanding Client/Customer Expectations
2. Methodology Transparency
3. The Use of Utilization Measures
4. The Appropriate Use of Adjustment to Achieve Comparison Group Equivalence