Development of Harmonized Outcome Measures for Use in Research and Clinical Practice

Michelle Leavy, MPH
OM1, Inc.

November 11, 2019
Funding

This work is supported by the Office of the Secretary Patient-Centered Outcomes Research Trust Fund under Interagency Agreement 18-596R-18 through Agency for Healthcare Research and Quality contract no. 75Q80119C00005.
Outcome Measures Framework

A standard, common model for patient and provider relevant outcome measures within and across condition areas
Variation in Outcome Definitions

Exacerbation Definitions Used in Asthma Registries

2017 GINA Report
Exacerbations of asthma are episodes characterized by a progressive increase in symptoms of shortness of breath, cough, wheezing or chest tightness and progressive decrease in lung function, i.e., they represent a change from the patient’s usual status that is sufficient to require a change in treatment.

NIH Workshop
An exacerbation is a worsening of asthma requiring the use of systemic corticosteroids (or for patients on a stable maintenance dose, an increase in the use of systemic corticosteroids) to prevent a serious outcome.

ATS/ERS Statement
Severe asthma exacerbations are events that require urgent action on the part of the patient and physician to prevent a serious outcome, such as hospitalization or death from asthma. Severe asthma exacerbations include at least one of the following:
(a) Use of systemic corticosteroids or an increase from a stable maintenance dose, for at least 3 days.
(b) A hospitalization or ER visit because of asthma, requiring systemic corticosteroids.

3 Reddel et al.. Am J Respir Crit Care Med. 2009 Jul 1;180(1):59-99
Why Is This So Hard?

- Different views on what constitutes an outcome measure
- Different goals in different studies
- Continuous reinventing of the wheel
- An industry that has grown up on quality and process measures
- The centrality of the patient not always considered
- No roadmap
- No organized way to harmonize differences
Outcome Measures Framework (OMF)

**Goal:** Common, conceptual model for classifying the range of outcomes that are relevant to patients and providers across most conditions

**Process:** Stakeholder-driven (~400) process incorporating iterative rounds of review and revision across multiple condition areas
Characteristics

**Participant**
- Demographics
- Genetics
- Family/Participant/Social History
- Functional/Performance Status
- Health Behaviors
- Environmental Exposures
- Preferences for Care

**Disease**
- Diagnosis
- Risk Factors
- Staging Systems
- Genetics of Disease
- Tissue or Infectious Agent
- Biomarkers
- Comorbidities/Symptoms
- Assessment Scales
- Physical Findings
- Severity
- Disease Understanding

**Provider**
- Training/Experience
- Geography
- Practice Setting
- Academic vs. Community

Treatment

**Type**
- Surgical
- Medical
- Device
- Alternative
- Education

**Intent**
- Palliative/Management vs. Curative

Outcomes

**Survival**
- Overall Mortality
- Cause-Specific Mortality
- Disease Free Survival
- Other

**Clinical Response**
- Recurrence/Exacerbation/
- Improvement/Progression/
- Change in Status/Other

**Events of Interest**
- Adverse Events/Exacerbations/
- Complications/
- Other

**Patient Reported**
- Functioning
- Quality of Life
- Other

**Resource Utilization**
- Inpatient Hospitalization/
- Office Visits/ED Visits/
- Productivity/
- Additional Treatments/
- Procedures/Direct Cost/Other

- Impact on Non-Participant
  Experience of Care

Building on Existing Efforts

And other efforts...
OMF Harmonization Project Goals

- Assess whether harmonized outcome measures can be developed for a sample set of 5 clinical areas:
  - Atrial fibrillation
  - Asthma
  - Depression
  - Lung cancer
  - Lumbar Spondylolisthesis

- Translate narrative harmonized definitions into standardized terminologies to facilitate consistent capture and extraction of measures from EHRs and registries

- Develop final report on policies and best practices for harmonization between registries and development of standardized libraries of outcome measures
Harmonization Using the OMF

An example from the Depression Workgroup
Methodology Overview

1. Recruited registries & stakeholders
2. Collected & categorized outcome measures using OMF
3. Built proposed minimum measure set
4. Harmonized definitions for measures in minimum set
5. Identified key characteristics to support risk adjustment
6. Produced final standardized library

Completed with 5 workgroup meetings over 8 months
## Participating Registries: Variations in Purposes, Patient Populations & Data

### Research
- UTSW Depression Cohort
- Dallas 2K
- Mood Network (PCORnet)
- NNDC Mood Outcomes Program
- Treatment-Resistant Depression

### Quality Improvement
- MN Community Measurement*
- PRIME Registry*
- PsychPRO*
- *Qualified Clinical Data Registry (QCDR)

### Health System / Population Management
- Dept. of Veterans Affairs
- Kaiser Permanente

---

**Narrow patient population, consistently collected detailed data**

**Broad patient population, variation in data consistency & detail**
## Participating Stakeholders

<table>
<thead>
<tr>
<th>Patient Advocacy Organizations</th>
<th>Professional Associations</th>
</tr>
</thead>
</table>
| • Depression and Bipolar Support Alliance  
• International Foundation for Research and Education on Depression  
• National Alliance on Mental Illness | • American Psychological Association  
• American Psychiatric Association  
• American Board of Family Medicine |

<table>
<thead>
<tr>
<th>Payers</th>
<th>Federal Agencies</th>
</tr>
</thead>
</table>
| • CMS  
• Blue Cross Blue Shield of Massachusetts | • FDA  
• National Institute of Mental Health  
• National Library of Medicine  
• SAMHSA  
• National Cancer Institute (PROMIS) |
Outcome Measures Collected from Registries & Other Sources*

- 27 outcomes categorized using the OMF
- The greatest number (n=11) were categorized as Clinical Response

*Other sources: ClinicalTrials.gov, World Health Organization, Peer-reviewed literature
**Examples of Submitted Measures**

- All-cause mortality
- Death from suicide
- Depression remission at 12 months
- Change in depressive symptoms
- Recurrence of depressive episode
- Suicide ideation and behavior
- Adverse events
- Functioning (physical, cognitive)
- Quality of life
- Change in social adjustment
- Depression-related resource utilization
- Depression-related hospitalization
# Depression Minimum Measure Set:

A minimum set of harmonized measures that can be captured consistently in research and clinical practice

<table>
<thead>
<tr>
<th>Survival</th>
<th>Events of Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-Cause Mortality</td>
<td>Adverse Events</td>
</tr>
<tr>
<td>Death from Suicide</td>
<td>(use of brief, publicly available validated measurement tool is recommended)</td>
</tr>
</tbody>
</table>

## Clinical Response

- Improvement in Depressive Symptoms:* Remission, Response
- Worsening in Depressive Symptoms:* Recurrence, Other**

*Timeframes
- 6 months (range = 4-8 months)
- 12 months (range = 10-14 months)

** Area for future investigation

## Events of Interest

- Suicide Ideation and Behavior (assessed via PHQ-9 for all patients; supplemental assessment for patients who indicate suicide ideation on PHQ-9)

## Patient Reported

- Depression-specific Quality of Life

## Resource Utilization

- Depression-related resource utilization
- Work productivity
Harmonization Process

- Compiled and compared detailed definitions of outcome measures in the minimum measure set to identify:
  - Measures for which more detailed definitions were needed to support harmonization
  - Measures that were distinct
  - Measures that addressed the same or similar concepts

- Through discussion with the workgroup, prioritized concept areas for harmonization

- Worked iteratively to harmonize outcome measure definitions
Step 1: Identified definitions for remission and response from registries, other sources
Harmonization Example: Remission & Response

**Step 2:**
Prepared detailed comparisons of definitions for workgroup discussion

### Depression Remission & Response Measures Comparison

<table>
<thead>
<tr>
<th>Quality Measures</th>
<th>Clinical Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Remission:</strong> Adult patients age 18 years and older with major depression or dysthymia and an initial PHQ-9 score &gt; 5 who demonstrate remission at six months defined as a PHQ-9 score less than 5.¹</td>
<td><strong>Remission:</strong> Remission was defined as an exit score of &lt; or = 7 on the 17-item Hamilton Depression Rating Scale (HAM-D) (primary outcome) or a score of &lt; or = 5 on the 16-item Quick Inventory of Depressive Symptomatology, Self-Report (QIDS-SR) (secondary outcome).³</td>
</tr>
<tr>
<td><strong>Response:</strong> Adult patients age 18 and older with major depression or dysthymia and an initial PHQ-9 greater than 9 who demonstrate a response to treatment at six months defined as a PHQ-9 score that is reduced by 50% or greater from the initial PHQ-9 score.²</td>
<td><strong>Response:</strong> Response was defined as ≥ 50% reduction in baseline 16-item Quick Inventory of Depressive Symptomatology-Self-Report (QIDS-SR-16) scores at exit.³</td>
</tr>
</tbody>
</table>

Remission and response measures designed for quality measurement rely on the PHQ-9.

Many registries use PHQ-9 as well.

But, clinical trials frequently use the HAM-D, MADRS, or QIDS-SR.

Crosswalks exist for:
- PROMIS Depression <-> PHQ-9
- QIDS-SR <-> HAM-D

---

² MN Community Measurement.
³ STAR*D Trial.
Harmonization Example: Remission & Response

Step 3:
Identified and discussed key differences in definitions, including review of validated instruments.
Harmonization Example: Remission & Response

**Step 4:** Arrived at recommended definition via workgroup discussions at in-person meeting, virtual meetings, & virtual activities

<table>
<thead>
<tr>
<th>Clinical Response</th>
<th>Improvement in Depressive Symptoms – Remission</th>
</tr>
</thead>
</table>
|                   | Patient age 18 or older with a diagnosis of major depression or dysthymia and an initial PHQ-9* score > 9 who demonstrates remission defined as a PHQ-9 score less than 5.  

*The PHQ-9 or another brief, publicly available, validated patient-reported instrument with empirically derived cutpoints equivalent to the PHQ-9 cutpoints for remission and response and for which an evidence-based crosswalk to the PHQ-9 exists should be used to measure clinical response. Other measures may be used in addition for research or other purposes.  

Timeframe for measurement:  
- 6 months (+/- 60 days)  
- 12 months (+/- 60 days)  

In some implementations, it would beneficial to capture earlier responses and remissions and to obtain higher degrees of follow-up. Additional measurements outside of the windows listed above are recommended as supplemental measures. |

<table>
<thead>
<tr>
<th>Clinical Response</th>
<th>Improvement in Depressive Symptoms – Response</th>
</tr>
</thead>
</table>
|                   | Patient age 18 or older with a diagnosis of major depression or dysthymia and an initial PHQ-9* score > 9 who demonstrates a response to treatment defined as a PHQ-9 score that is reduced by 50% or greater from the initial PHQ-9 score.  

*The PHQ-9 or another brief, publicly available, validated patient-reported instrument with empirically derived cutpoints equivalent to the PHQ-9 cutpoints for remission and response and for which an evidence-based crosswalk to the PHQ-9 exists should be used to measure clinical response. Other measures may be used in addition for research or other purposes.  

Timeframe for measurement:  
- 6 months (+/- 60 days)  
- 12 months (+/- 60 days)  

In some implementations, it would beneficial to capture earlier responses and remissions and to obtain higher degrees of follow-up. Additional measurements outside of the windows listed above are recommended as supplemental measures. |
Translation to Standardized Terminologies

• Narrative definitions were mapped to standardized terminologies

• For each outcome, the following were defined:
  o An object representing the outcome condition itself: In many cases, the only structured data will be an assertion of an outcome, with all the supporting evidence being present in the narrative
  o FHIR resources for evidence for the outcome: These include labs, diagnostic imaging, etc.
  o FHIR resources for additional relevant events: These might include procedures, encounters, etc.
  o Temporal aspects for all events: These allow for inferred relationships
To build connections across initiatives, the following sources were searched for overlap:

- **eCQI Resource Center**: Primarily looking for overlapping criteria
- **Value Set Authority Center (VSAC)**: Primarily looking for overlapping value sets
- **C-CDA**: Primarily looking for overlapping data representations
- **NIH Common Data Element (CDE) Resource Portal**: Primarily looking for overlapping data element definitions
## Example of Standardized Definition

### Appendix B. Standardized Library of Depression

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Definition</th>
<th>Population Criteria - Data Objects and Expression Logic: Initial Population (or Modified for Individual Outcome)</th>
<th>Population Criteria - Data Objects and Expression Logic: Outcome Focused Population</th>
<th>Data Criteria / Value Sets</th>
</tr>
</thead>
</table>
| Death from suicide                                | Patient with a diagnosis of major depression or dysthymia who died from suicide, reported in 12-month intervals.                                                                                           | o Condition  
  - code: in "Major Depression Condition" or "Dysthymia Condition" value set  
  - onset(x)  
  - clinicalStatus: any (active/inactive/remissed/resolved, etc)  | o AND: Condition  
  - code: in "Expired" value set  
  - Condition: dueTo extension  
  - code: in "Suicide Condition" value set  
  - onsetDateTime: during parameter "12-month interval of interest"  | • "Major Depression" value set  
• "Dysthymia" value set  
• "Expired" value set  
• "Suicide Condition" value set |
| Improvement in Depressive Symptoms – Remission    | Patient age 18 or older with a diagnosis of major depression or dysthymia and an initial PHQ-9* score > 9 who demonstrates remission defined as a PHQ-9 score less than 5.  
*The PHQ-9 or another brief, publicly available, validated patient-reported instrument with empirically derived cutpoints equivalent to the PHQ-9 cutpoints for remission and response and for which an evidence-based crosswalk to the PHQ-9 exists should be used to measure clinical response. Other measures may be used in addition for research or other purposes. | o Condition  
  - code: in "Major Depression Condition" or "Dysthymia Condition" value set  
  - onset(x)  
  - clinicalStatus: any (active/inactive/remissed/resolved, etc)  | o AND: Observation  
  - Code: final  
  - Code: "PHQ-9 effectiveDate" value | • "Major Depression Condition" value set  
• "Major Depression ICD10 set" |
|                                                   |                                                                                                                                                                                                           |                                                                                                                           | • valueQuantity  
  - PHQ-9 observation  
  - PHQ-9 observation | • ICD10 / SNOMED  
• ICD-10-SC  
• ICD-10-CM  
• F320  
• F321  
• F322  
• F323  
• F324  
• F325  
• F329  
• F330  
• F331  
• F332  
• F333  
• F334  
• F340  
• F341  
• Major depressive disorder, single episode, mild  
• Major depressive disorder, single episode, moderate  
• Major depressive disorder, single episode, severe without psychotic features  
• Major depressive disorder, single episode, severe with psychotic features  
• Major depressive disorder, single episode, in partial remission  
• Major depressive disorder, single episode, in full remission  
• Major depressive disorder, single episode, unspecified  
• Major depressive disorder, recurrent, mild  
• Major depressive disorder, recurrent, moderate  
• Major depressive disorder, recurrent severe without psychotic features  
• Major depressive disorder, recurrent, severe with psychotic symptoms  
• Major depressive disorder, recurrent, in remission, unspecified  
• Major depressive disorder, recurrent, in partial remission  |
Final Products

- Minimum measure sets and standardized libraries for 5 condition areas
- Final Report on methods, lessons learned (available on AHRQ website)
- Publications:
  - Atrial fib – published in heart Rhythm in January 2019
  - Asthma – published in Journal of Allergy and Clinical Immunology in September 2019
  - 4 publications in process

---

### Appendix B. Standardized Library of Depression

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Definition</th>
<th>Population Criteria - Data Objects and Expression Logic: MHDR (Where applicable)</th>
<th>Population Criteria - Data Objects and Expression Logic</th>
<th>Focused Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND</td>
<td>Condition</td>
<td>• code: in &quot;Expired&quot; value set</td>
<td>Condition: dueTo-observation</td>
<td>• code: in &quot;Suicide Condition&quot; value set</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ND</td>
</tr>
</tbody>
</table>

### Final Report

Development of Harmonized Outcome Measures for Use in Patient Registries and Clinical Practice: Methods and Lessons Learned

Prepared for:
Agency for Healthcare Res.
U.S. Department of Health
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. 296-2014-16

Prepared by:
L&M Policy Research, LLC
Washington, DC 20006

With partners: OM1, AcademyHealth

**Investigators:**
Michelle B. Leavy, MPH
Claudia Sehr, PhD
Felicitas G. Kansamal, MS,
Margaret Elder Johnson, RN
Raj Sabharwal, MPH
Paul Wallace, MD
Richard E. Gliklich, MD
1 OM1, Inc.
2 L&M Policy Research, LLC
3 AcademyHealth

**AHRQ Publication No. 1**
February 2019

**NEWS FROM THE HEART RHYTHM SOCIETY**

**Harmonized outcome measures for use in atrial fibrillation patient registries and clinical practice**

**Endorsed by the Heart Rhythm Society Board of Trustees**

Hugh Callans, MD, FHRS, 
Richard E. Gliklich, MD, MPH
Jonathan F. Piccirilli, MD, MHS, FHRS
Jonathan C. Hou, MD, MHS, FHRS
Sanghavi Arunachalam, MD, MS, FHRS
William Lewis, MD, FHRS
Saman Nazarian, MD, PhD, FHRS
**Witu P. Batalushe, MD, MSc, FAACC, FHRS**

From the Cardiology Division, Johns Hopkins University School of Medicine, Philadelphia, Pennsylvania, 1411 Park Road, Stanford, California, and Harward Medical School.

**BACKGROUND:** Atrial fibrillation (AF) affects an estimated 2.5 million people worldwide, leading to increased morbidity and increased risk of heart failure and stroke. Many AF patients exist, but the ability to link and compare data across registries and a lack of harmonization.

**OBJECTIVES:** The purpose of this project was to develop a set of harmonized outcomes measures that could be used in patient registries and clinical practice.

**METHODS:** At patient registries were identified through searches and linked to the workgroup and submitted for review. Additional measures were identified through searches and review of consensus statements. Outcome were categorized using the Agency for Healthcare Research and Quality’s (AHRQ) supported Outcome Measures Framework (OMF). 17

---

### Harmonized Outcome Measures for Use in Asthma Patient Registries and Clinical Practice

**Richard E. Gliklich MD, MPH1, 2; A. Carrol M.D., MPP, MSCI, 3; Michelle B. Leavy MPH, MSCI, 2; Valerie G. Press MD, MPH, 1; Archara Bunawan KBBS, MHRM, 4; Christopher L. Carroll MD, MSCI, 3; Julian Harris MBA, 4; Sarah S. Bitran PhD, 5; Robert Frutchot MD, MPH 1; Reynald A. Panettieri Jr. MD, PhD 6; Guiselle S. Mascanzi MD, MSCI 1**

**Background:** Asthma, a chronic airway disorder, affects an estimated 25 million people in the United States and 330 million people worldwide. Although many asthma patient registries exist, the ability to link and compare data across registries is hindered by a lack of harmonization in the outcome measures collected by each registry.

**Objectives:** The purpose of this project was to develop a minimum set of patient and provider...
Lessons Learned & Next Steps
Lessons Learned

• Harmonization is feasible
  • Each condition area presented unique challenges, but the OMF provided an effective model to guide discussions across conditions

• Broad stakeholder and registry support
  • High participation rates in registry workgroups, with major national registries in each condition area represented

• More work is needed to support routine, consistent collection of PROs
  • More work is needed to identify or develop PROs that are patient and clinician-relevant and can be captured on a routine basis in asthma, lung cancer, and atrial fibrillation
  • More work is needed to understand burden of capture and appropriate intervals for measurement in depression and lumbar spondylolisthesis
Possible Barriers to Implementation

Stakeholders in the prior project expressed enthusiasm for the standardized outcome measures, but they identified several potential barriers:

- How difficult will it be to extract the necessary data from different EHRs?
- Is capture of the PROs on a regular basis too burdensome for patients and clinicians?
- Are the measures useful for informing clinical decision-making and understanding and improving patient outcomes?
- Will capture of the harmonized outcome measures improve the utility of the registry data for conducting patient-centered outcomes research?
Title: Outcome Measure Harmonization and Data Infrastructure for Patient Centered Outcomes Research in Depression

Purpose: To demonstrate that capturing the harmonized depression outcome measures in the clinical workflow and submitting these data to different registries can:
- Improve clinical care
- Reduce the burden of registry participation
- Increase the utility of registry data for research purposes
Depression Pilot Project Objectives

1. Demonstrate that collection of the harmonized depression outcome measures is feasible, sustainable, and useful for clinicians participating in primary care and mental health patient registries

2. Demonstrate that collection of the harmonized outcome measures is feasible, sustainable, and useful for clinicians in a health system setting

3. Evaluate whether collection of the harmonized measures increases the utility of registry data for research purposes
# PRIME Registry

<table>
<thead>
<tr>
<th>Sponsor</th>
<th>American Board of Family Medicine (ABFM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Outpatient quality improvement registry open to all primary care clinicians. The registry is a Qualified Clinical Data Registry (QCDR), enabling it to produce and test measures that can be certified by CMS for the Quality Payment Program independent of other measure approval processes</td>
</tr>
</tbody>
</table>
| Purposes                 | • Capture data from EHRs to calculate quality measures  
• Provide a dashboard to enable clinicians to track and measure performance data at the practice, clinician, and patient level  
• Provide burden-reducing tools for reporting and practice improvement  
• Provide data for research purposes  
• Provide social determinant information and connection to local resources in relation to clusters of disease and poor outcomes in patient population |
| Participating Sites      | Over 900 sites                           |
### PsychPRO

<table>
<thead>
<tr>
<th>Sponsor</th>
<th>American Psychiatric Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Outpatient quality improvement registry open to mental health professionals. The registry is a Qualified Clinical Data Registry (QCDR).</td>
</tr>
</tbody>
</table>
| Purposes      | • Help psychiatrists and other mental health professionals:  
  • Meet Medicare quality reporting requirements  
  • Maintain professional recertification  
  • Access clinical decision-support tools to inform evidence-based  
  • Provide data for research purposes |
Selected Outcome Measures

1. Improvement in Depressive Symptoms – Response
2. Improvement in Depressive Symptoms – Remission
3. Worsening in Depressive Symptoms – Recurrence
4. Adverse Events (related to depression treatment)
5. Suicide Ideation and Behavior
6. Death from Suicide

Measured using the PHQ-9
Project Overview

**Workstream 1**
Implement Measures in Registries

**Objective:** Demonstrate that collection of the harmonized outcome measures is feasible, sustainable, and useful for clinicians participating in primary care and mental health patient registries.

**Tasks:**
- Implement 6 outcome measures in PsychPRO and PRIME Registry
- Capture PHQ-9 data at baseline, 6 and 12 months at subset of registry sites
- Provide measure results back to clinicians in registry portal
- Assess burden and value from registry perspective

---

**Workstream 2**
Implement Measures in Health System

**Objective:** Demonstrate that collection of the harmonized outcome measures is feasible, sustainable, and useful for clinicians in a health system setting.

**Tasks:**
- Using an open source app, capture selected outcome measures, including PHQ-9 information, and provide results in clinician’s workflow
- Assess burden of capturing the measures from patients (PHQ-9) and clinicians and value of providing the measure results

---

**Workstream 3**
Research Demonstration

**Objective:** Evaluate whether collection of the harmonized measures increased the utility of registry data for research purposes.

**Tasks:**
- Evaluate utility of using natural language processing (NLP) to extract relevant information from unstructured EHR data (proof of concept)
- Conduct pilot data analysis using combined de-identified data from PRIME Registry and PsychPRO
- Develop governance & data use agreement templates

---

**STAKEHOLDER PANEL:**
Provide feedback & guidance across each workstream
Thank You!

For More Information:

Michelle Leavy
OM1
mleavy@om1.com