DATA FOR IMPROVEMENT
MODULE #2: DATA QUALITY WORKSHEET

CALIFORNIA ASSOCIATION of PUBLIC HOSPITALS AND HEALTH SYSTEMS

SNI CALIFORNIA HEALTH CARE SAFETY NET INSTITUTE

SYSTEM:

TEAM MEMBER:

PRIME MEASURE:

Visit SNI Link for additional resources

INSTRUCTIONS: For this homework, you should identify a PRIME measure that figures centrally to your PRIME program and includes at least some data from source systems that your organization has authority over. You may also suspect that there are quality deficiencies in the numerator and denominator. This worksheet should help your team identify some of the data quality issues associated with the PRIME measure you have chosen. **Use the "your findings" sections of the worksheet to answer all questions that apply to the measure you have chosen. If the question does not apply, you may skip it and continue on to the next question. Place a checkmark in the box of each question you answer.**

SUBMIT HOMEWORK BY 6/30 TO: RONLI LEVI (RLEVI@CAPH.ORG)

Part 1: Identifying Immediate Issues

Questions to Consider

Issue Impact	☐1. What are the specific, immediate symptoms that you have noticed regarding this measure (i.e. "numerators/denominators are inconsistent depending on the system used to run the report")?
	☐ 2. How do these symptoms impact reporting on the measure?
	☐ 3. What is the impact of the issue more generally? What processes are being used to work around it?
Stakeholder Input	☐ 1. Who are the teams associated with the measure in the data collection process?
	☐ 2. Who are the external stakeholders associated with this process?
	☐ 3. Have you established a pathway to easily investigate workflows among all of these teams? If so, please describe.
YOUR FINDINGS:	

Part 2: Create a Problem Statement

Problem Statement	☐ 1. What is your problem statement for this measure? Include the following: • **Description of problem** (e.g. the observed irregularities in a given measure) • **Relevant data pathway and chain of stewardship** • **Impact of problem** (how does the issue impact business, clinical, and operational processes and outcomes) • **Existing approach** to addressing problem, if any (e.g. current workarounds) • **Hypothesis** regarding the cause of data quality problem
YOUR PROBLEM STATEMENT	

Part 3: Analyzing Data Integrity

Consistency	☐ 1. Do you receive the same measured results from different sources (e.g. EHR registry vs. PHM software)?
	☐ 2. When charting a sample from reported data, do numbers in patient records match report?
	☐ 3. Are patient records de-duplicated?
	☐ 4. Is there a referenceable audit trail for the data that can be leveraged?
	☐ 5. Are the correct codes consistently associated with the test or diagnosis being measured?
	☐ 6. Is there a referenceable audit trail for the data that can be leveraged?
YOUR FINDINGS	

Reliability	
	☐ 1. Are data entered into proper EHR fields?
	☐ 2. Are EHR fields free of text elements that could nullify data (e.g. text in numeric field)?
	☐ 3. Have any elements of the data been transposed or mapped?
	☐ 4. Are data recorded properly (e.g. BP readings being rounded up or down before entry)?
	☐ 5. Are data appropriately limited to only cover sites, providers, and patients that are exposed to the QI interventions?
	☐ 6. Are all eligible data elements being captured in the report logic (e.g. relevant ICD codes)?
	☐ 7. Is the date of a procedure based on the order date or the completion date? (If by order date could show false positives if procedure not completed)
	☐ 8. Do start date/times occur before stop date/times reliably?
	☐ 9. If the measure contains two linked data-points, is there enforcement for data being present among both variables, and/or are illogical entries disallowed (e.g. being both "pregnant" AND "male")?
YOUR FINDINGS	
Accuracy	☐ 1. Are numerators/denominators relatively consistent over time? Is there a logical reason if not?
	☐ 2. Are tests being counted as completed when there is text in a results field? If so, are results always entered or is other text entered in the field that may falsely indicate completion?
	☐ 3. Is the report logic querying data from the proper EHR fields?
	☐ 4. Are patients who "age out" of the measure specifications or who develop an exclusion properly excluded at the appropriate time?
	☐ 5. Have pertinent codes (LOINC or ICD) changed? (e.g. new ones added or old ones retired)
	☐ 6. Is there a logical relationship between proportions? (e.g. diagnosed HTN cohort numerators should add up to the diagnosed cohort denominator)
	☐ 7. Are results very similar for reports one day of the week vs. another?

YOUR FINDINGS	
Completeness	☐ 1. Are elements consistently documented in structured fields or with standardized free text?
	☐ 2. Are scanned patient documents/reports attached to orders so that they can be queried?
	☐ 3. Are critical data points (e.g. BP readings) documented in a telephone encounter or other place in the EHR that cannot be queried?
	☐ 4. If a repeat measure is taken, is it entered into the reportable vitals field, or into free text (where it might be retrieved by the measure query)?
	☐ 5. Do start dates and end dates align with the measure specifications?
	☐ 6. Are select sites, providers, or patients missing from the data set?
YOUR FINDINGS	

Part 4: Preparing to Remediate

Data Quality Determination	☐ 1. What are the next steps that will be taken to develop a remediation plan?
YOUR FINDINGS	

Submit homework by June 30th to: Ronli Levi (rlevi@caph.org)