

WHAT TO LOOK FOR IN EVALUATIONS

EVALUATION CHARACTERISTICS	MOST RIGOROUS...	...LESS RIGOROUS BUT STILL INFORMATIVE	...AT MINIMUM SHOULD INCLUDE
Is Objective	Completed by neutral third party	Data collected by external group but analyzed in-house	In-house data monitoring and analysis
Measures a “Baseline”	Assesses participants and control/comparison group along critical metrics (e.g., prevalence of malaria and anemia) before beginning interventions	May use rapid assessment techniques in a small subset area to get a general sense of the current malaria burden	Considers available regional data that can be used as a benchmark
Provides a Comparison or Control Group	Randomly assigns a portion of eligible children or communities to a control group or randomly staggers introduction of a new model or tool; Ensures the difference between the groups is not larger than what chance would create	Carefully matches communities with comparable ones on key characteristics or uses statistical techniques to “control for” differences at project’s start	Considers externally calculated national, district and/or other comparable measures
Includes a Sufficient Number of Individuals or Communities	Sample size depends on the size of effect anticipated from the program; the larger the effect size, the fewer communities are required for it to be found “statistically significant.” Even a small pilot of a dozen communities might be sufficient to inform practitioners on how to improve practice and whether the program is worth scaling up		
Uses Objective and Meaningful Measures	Measures actual health impacts (e.g., child mortality) Uses both quantitative and qualitative methods Employs surveys and indicators that have been externally assessed for their ability to measure the intended factor consistently across the population of interest	Measures change in coverage or use of key health interventions (e.g., bednet use) during project cycle using standardized household surveys Presents reports from multiple stakeholders (e.g. health staff, beneficiary communities) and identifies likely biases	Monitors project outputs (e.g., medications distributed) and quality of services Considers reports from one party and identifies likely biases
Considers Program’s Replicability	Evaluates multiple implementations in a diversity of sites	Evaluates multiple implementations in similar settings	Evaluates a single implementation