STAR RATING SYSTEM AND EVALUATION

A MIXED METHODS ANALYSIS OF TANZANIA'S NATIONAL HEALTH FACILITY QUALITY ASSESSMENT SYSTEM

SANAM RODER-DEWAN, MD, DRPH
TALHIYA YAHYA, MD, MPH
Principal Investigator: Sanam Roder-DeWan, MD, DrPH (UNICEF Tanzania)

Co-principal Investigator: Donat Shamba, MSc (Ifakara Health Institute)

Research Scientists
- Anna Gage, MSc (Harvard Chan School of Public Health)
- Mary Ramesh, MPH (Ifakara Health Institute)
- Magreat Somba, MA (Ifakara Health Institute)

Policy and program advisors
- Eliudi Eliakimu, MD, MPH (MOH)
- Talhiya Yahya, MD, MPH (MOH)
- Mohammed Mohammed, MD, MPH, MSc (World Bank)

Study coordinator
- Heller Rajab, BAC (Ifakara Health Institute)
STAR RATING SYSTEM
TALHIYA YAHYA, MD, MPH
Ministry of Health Initiative started in 2014 during Big Results Now Initiative
System was designed to be aligned with key health system initiatives
Four workstreams
- Human resources
- Commodities
- RMNCH
- Performance management
Conducted twice to-date (2015/16 and 2017/18)
**Initiative**

Assess, rate and develop specific facility improvement plans for primary level health facilities

**Approach**

- Develop a specialised health facilities assessment tool based on the identified issues on the ground
- Roll-out the assessment tool in all primary health facilities in all regions
- While assessing the health facilities, specific gaps are identified for improvement
- Health facilities are rated after assessment
- Specific intervention programmes are rolled-out at selected regions to improve the health facilities below 3-Star

**Aspiration**

80% of primary health facilities in selected regions will be 3-Star and above by 2018

---

**6760 health facilities are not holistically assessed**

- Arusha
- Dodoma
- DSM
- Geita
- Iringa
- Kagera
- Katavi
- Kigoma
- Kilimanjaro
- Lindi
- Manyara
- Mara
- Mbeya
- Morogoro
- Mtwar
- Mwanza
- Njombe
- Pwani
- Rukwa
- Ruvuma
- Shinyanga
- Simiyu
- Singida
- Tabora
- Tanga

Source(s): BRN Healthcare (2014)
DEVELOPING THE TOOL

- Participatory process
- Comparative studies were conducted looking at other assessment tools
  - SafeCare
  - Results-Based Financing
  - Pay-for-Performance
  - Service Provision Assessment
- Key question – use existing tools or create our own

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Drawbacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using existing measuring tools</td>
<td>Create own measuring tools in BRN lab</td>
</tr>
<tr>
<td>- Ready-to-use</td>
<td>- Full ownership by MoHSW</td>
</tr>
<tr>
<td>- Internationally accredited (SafeCare)</td>
<td>- Data is owned by the government</td>
</tr>
<tr>
<td></td>
<td>- High impact – addressing on-the-ground issues</td>
</tr>
<tr>
<td></td>
<td>- Does not address HRH performance issues</td>
</tr>
<tr>
<td></td>
<td>- Need expertise from a person who has widespread experience in the healthcare and criteria development</td>
</tr>
<tr>
<td></td>
<td>- Requires funds</td>
</tr>
<tr>
<td></td>
<td>- The tool needs to be developed and tested</td>
</tr>
<tr>
<td></td>
<td>- Requires fully-trained assessors</td>
</tr>
</tbody>
</table>
## ASSESSMENT AREAS

<table>
<thead>
<tr>
<th>Assessment Area</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Health Facility Management</td>
<td>12 indicators</td>
</tr>
<tr>
<td>3. Use of facility data for service improvements</td>
<td>6 indicators</td>
</tr>
<tr>
<td>4. Staff Performance Management</td>
<td>5 indicators</td>
</tr>
<tr>
<td>5. Organisation of services</td>
<td>8 indicators</td>
</tr>
<tr>
<td>6. Handling of emergencies/referral</td>
<td>7 indicators</td>
</tr>
<tr>
<td>7. Client Focus</td>
<td>4 indicators</td>
</tr>
<tr>
<td>8. Social accountability</td>
<td>7 indicators</td>
</tr>
<tr>
<td>9. Facility infrastructure</td>
<td>14 indicators</td>
</tr>
<tr>
<td>10. Infection Prevention and Control</td>
<td>11 indicators</td>
</tr>
<tr>
<td>11. Clinical Services</td>
<td>13 indicators</td>
</tr>
<tr>
<td>12. Clinical Support Services</td>
<td>20 indicators</td>
</tr>
<tr>
<td>Minimum Score in Four Domains</td>
<td>0-Star</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>0-19%</td>
<td>20-39%</td>
</tr>
</tbody>
</table>
1. MoHSW (DCS, DPS, DHR)
2. Muhimbili National Hospital
3. National Institute for Medical Research
4. Medical Store Department
5. Tanzania Food and Drug Authority

**National Team (HSIQAS)**
(to cover 24 regions and 3 municipals of Dar es Salaam Region)

**Regional Team**
(6 per council)

**Council Team**
6 assessors/LGA
(assessors to jointly assess district hospitals)

**MULTI-LEVEL ASSESSMENT PROCESS**

- Team A: 2 assessors/disp & HCs
- Team B: 2 assessors/disp & HCs
- Team C: 2 assessors/disp & HCs
<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of district hospitals, health centers and dispensaries</td>
<td>6996</td>
</tr>
<tr>
<td>Total number of assessors</td>
<td>1140</td>
</tr>
<tr>
<td>Estimated assessment duration per region</td>
<td>~4 weeks</td>
</tr>
<tr>
<td>Estimated assessment cost</td>
<td>$200 per facility</td>
</tr>
</tbody>
</table>
EVALUATION FINDINGS
RESEARCH QUESTIONS

1. What factors are associated with variation in facility capacity to improve?
2. What are the mechanisms through which facility-level quality improvement occurs using the Star Rating System?
3. How does the Star Rating instrument perform as a quality measurement tool?
A high quality health system is one that optimizes health care in a given context by consistently delivering care that improves or maintains health outcomes, by being valued and trusted by all people, and by responding to changing population needs.
DATA SOURCES AND QUALITY CONSTRUCTS

Clinical observations
Provider surveys
Provider interviews
Facility audits

Patient pre-visit interviews
Patient exit interviews

Facility audit
Provider surveys
Provider interviews

6 data collection modalities
QUANTITATIVE DATA

1. **Star Rating Data Set**
   - 2 rounds of data collection (2015/16 and 2017/18)
   - All primary care facilities on the mainland excluding Dar es Salaam
   - N=5595 facilities

2. **Evaluation Data Set**
   - All regions on mainland
   - 47 districts
   - 280 facilities
     - 609 provider interviews
     - 1,275 client interviews and observation

3. **Additional survey data**
   - Service provision assessments 2016
   - DHS 2015
   - World Pop 2015
   - Geographic data
     - Open street map
     - Natural earth
QUANTITATIVE EVALUATION SAMPLE

- 47 districts
- 3 dispensaries
- 3 health centers
- 3 providers
- 5 clients
**QUALITATIVE EVALUATION SAMPLE**

<table>
<thead>
<tr>
<th>Most improved</th>
<th>Least improved</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Southern Highland</strong></td>
<td><strong>Southern Highland</strong></td>
</tr>
<tr>
<td><strong>Lake Zone</strong></td>
<td><strong>Lake Zone</strong></td>
</tr>
<tr>
<td><strong>Dispensary</strong></td>
<td><strong>Dispensary</strong></td>
</tr>
<tr>
<td><strong>Health center</strong></td>
<td><strong>Health center</strong></td>
</tr>
<tr>
<td>N=3</td>
<td>N=4</td>
</tr>
<tr>
<td>N=4</td>
<td>N=3</td>
</tr>
<tr>
<td>N=3</td>
<td>N=4</td>
</tr>
<tr>
<td>N=4</td>
<td>N=3</td>
</tr>
</tbody>
</table>

- Most improved = 2 or 3 star change
- Least improved = 0 or -1 star change

N=27
MIXED METHODS

- Quantitative
  - Descriptive statistics
  - Multivariable regression
  - Geospatial methods
- Qualitative
  - Thematic content analysis of interviews using grounded theory
  - Open coding until saturation of codes
- Mixed
  - Simultaneous collection and convergent analysis
  - Multi-disciplinary discussion of results → local theory of improvement

Diagram:
- Read
- Code
- Organize
- Purge
- Consolidate

Flow:
- Most improved
- Least improved
RESEARCHER BIAS

- Facilities know their weaknesses
- Changes to system are primary drivers of improvement: funding, workforce
- Participation vs. top-down programming
WHAT FACTORS ARE ASSOCIATED WITH VARIATION IN FACILITY CAPACITY TO IMPROVE?

FINDINGS: QUESTION 1
### STAR RATING ASSESSMENT RESULTS

- **72% of facilities improved**
- **27% improved by 2+ stars**
- **45% improved by one star**
- **28% same or lower score**

<table>
<thead>
<tr>
<th>Mean Star Rating scores in country, N=5807</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Star rating (0-5)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Overall score (0-100)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Domain A. Facility management and staff performance</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Domain B. Fulfilment of service charters and accountability</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Domain C. Safe and conducive facilities</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Domain D. Quality of care and services</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
• No difference in facility type
• Public facilities improved more than private for-profit and non-profit facilities, but also had lower baseline
• Strong baseline effect

N=5595
GEOGRAPHIC FACTORS – A LOCAL EFFECT

- Greatest improvements (averages up to 1.75 stars)
  - Lake Zone (except Mara)
  - Pwani region
- Least improvement
  - Southern Highlands Zone (except Mbeya)
- 11% of the total variation in ability to improve was due to differences between regions
- 14% was due to differences between districts within regions
- Beyond administrative boundaries, there is a local neighborhood effect where facilities near improving facilities also improved
POTENTIAL IMPLEMENTATION FACTORS

Outer setting
- Patient needs and resources
- Cosmopolitanism
- Structural environment
- Peer pressure
- External policies and incentives

Inner setting
- Structural characteristics
- Networks and communication
- Culture

Council administration

Space/Proximity

Facility improvement
### IMPLEMENTATION FACTORS ASSOCIATED WITH IMPROVEMENT

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coef.</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of people within 5 km radius of facility (ln)</td>
<td>0.8</td>
<td>(0.4,1.2)</td>
</tr>
<tr>
<td>Institutional delivery percent in council</td>
<td>1.7</td>
<td>(-6.1,9.5)</td>
</tr>
<tr>
<td>Female primary education percent in council</td>
<td>3.8</td>
<td>(-8.1,5.6)</td>
</tr>
<tr>
<td>Healthcare decisions percent in council</td>
<td>-8.7</td>
<td>(-22.9,5.6)</td>
</tr>
<tr>
<td>Number of facilities in council</td>
<td>0.0</td>
<td>(-0.1,0)</td>
</tr>
<tr>
<td>Distance to major road (ln)</td>
<td>-0.1</td>
<td>(-0.2,0.1)</td>
</tr>
<tr>
<td>Distance to large city (10 kms)</td>
<td>0.1</td>
<td>(-0.1,0.2)</td>
</tr>
<tr>
<td>Percentile rank at baseline</td>
<td>3.6</td>
<td>(0.5,6.7)</td>
</tr>
<tr>
<td>RBF Participation</td>
<td>6.0</td>
<td>(4.3,7.6)</td>
</tr>
<tr>
<td>RBF ineligibility due to low baseline</td>
<td>11.0</td>
<td>(9.2,12.9)</td>
</tr>
<tr>
<td>Ownership (Public ref.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>-4.2</td>
<td>(-5.5,-2.9)</td>
</tr>
<tr>
<td>Non-profit</td>
<td>-2.1</td>
<td>(-3.3,-1)</td>
</tr>
<tr>
<td>Level (Dispensary ref.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health center</td>
<td>3.0</td>
<td>(2.4)</td>
</tr>
<tr>
<td>Primary hospital</td>
<td>7.3</td>
<td>(5.4,9.3)</td>
</tr>
<tr>
<td>Baseline performance</td>
<td>-10.0</td>
<td>(-11.6,-8.4)</td>
</tr>
</tbody>
</table>

- Multivariable regression using change in Star Rating Score (0-100) as the outcome
- Population density associated with improvement
- Facilities with RBF improved by 6 points more than non-RBF facilities, but facilities that were ineligible for RBF because they scored zero at baseline improved by 11 points
- Public facilities improved more than private for- or non-profit facilities
- Hospitals and health centers improved more than dispensaries
- Strong baseline effect: most of the improvement came from very low performing facilities
WHAT ARE THE MECHANISMS THROUGH WHICH FACILITY-LEVEL QUALITY IMPROVEMENT OCCURS USING THE STAR RATING SYSTEM?

FINDINGS: QUESTION 2
THREE CORE MECHANISMS

1. Accountability
2. Learning
3. Group identity and benchmarking
MECHANISM 1: ACCOUNTABILITY

- Internal accountability
  - Having a common language
  - Data
  - Hierarchy is important - Coming from MOH
  - CHMT involvement

- External accountability
  - Health Facility Governing Board
  - Suggestion boxes (mixed results)
  - “Sweet language”
  - Quality->utilization->funds
MECHANISM 2: LEARNING

- Respondents learn from the assessment process
  - Learn about weaknesses
  - Learn about how to deliver high quality care
  - Stay current
  - Make up for gaps in pre-service education

- …not everything I deal with in the facility I was taught in school. Other things I was not taught…When assessing they will find weakness whereby in my point of view I thought it was a normal thing but them they see [no]. (Ibaba)

- At first, as I told you, we didn’t know a lot of things but after they came we knew we needed to have a lot of things. Like when serving a customer we never paid attention to how long to we need to attend to one customer but after the assessment we were told it was important to use time when attending a customer, so per day we watch the average we used to serve how many clients. (Kisale)

- Another thing that I thought was good about star rating was how they used guidelines to inspect, there were things that I realize that day I was suppose to have just by the question they ask. (Ngulugulu)
Because the aim is to emphasize people, especially us the provider in the facilities that we have a lower rank, we should feel jealous. Why should your fellow get 2 stars while you have one, or why he has 3 you have 2? (Mwengemshindo)

...being one of the providers in the facility you are among the reasons of the facility to perform well or you might be among the reasons for the facility’s down fall. (Bweri)

With BRN you know what you are supposed to do, the system has inspired providers, even creating competition among facilities. (Tenende)

So they will be the one to celebrate, they even ate pilau….they danced, they had a lot to drink, so it helped them. (Bweri)
QUALITY ACTIONS

I. Advocacy

- For instance, in 2017/2018 it was seen that we don’t have administrative building and the Matron office is in the ward. It was suggested that, the council together with the region will take this information to the Ministry if possible we should get the administrative building also paediatric ward should be improved. (Murangi)

II. Self-checking

- ...so you will just know, that day they came they asked these questions, so when you are about to do something that involves that question asked you will do it with caution since you know you are doing it to help improve star ratings. (Bweri)

III. Collaboration

- The first thing was collaboration of providers because if there was no collaboration then none of these could have been possible. So what happened was after getting that report we sat all the providers and noted all that was written in the work plan (Kisale)
MOST - LEAST

- **Divergent**
  - Bureaucracy
  - No ownership over improvement
  - Supervision not happening
  - Don’t know about stars or baseline

- **Convergent**
  - No control over human resources
  - Funding as an enabler
LOCAL THEORY OF IMPROVEMENT

SRA

Accountability
Learning
Group identity/benchmarking
Advocacy
Self checking
Collaboration

Inputs (Funds, HR, Equipment)

Improved quality processes
Pay-for-performance
Happy providers

Health insurance and fee-for-service

Increased utilization
HOW DOES THE STAR RATING INSTRUMENT PERFORM AS A QUALITY MEASUREMENT TOOL?

FINDINGS: QUESTION 3
How does the star rating map to other quality measures?

- Good medical practice
- User experience
- Processes of care
  - Competent care and systems
  - Positive user experience
- Quality impacts
  - Better health
  - Confidence in system
  - Economic benefit
- Learning and improvement
- Foundations
  - Population health needs and expectations
  - Governance policy, insurance, and non-health sectors
  - Platforms accessibility and organisation of care
  - Workforce numbers, skill, and support
  - Tools equipment, medicines, and data
- Basic infrastructure and equipment
- Management index
- Diagnostic and treatment accuracy
HOW DOES THE STAR RATING MAP TO OTHER QUALITY MEASURES?

- Quality measures increase with each level of star rating except for user experience.
- Star rating does not reflect user experience.
- Star rating only discriminates diagnostic and treatment accuracy at low level of performance.
- Management index has largest discriminatory power.

Mean facility performance by Star Rating score at reassessment

- Basic infrastructure and equipment
- User experience
- Management index
- Good medical practice
- Diagnosis and treatment accuracy

Star Rating Score

Quality index score

Mean facility performance by Star Rating score at reassessment:

- 0-star (N=7)
- 1-star (N=72)
- 2-star (N=118)
- 3-star (N=71)
- 4-star (N=12)
HOW DOES THE STAR RATING MAP TO OTHER QUALITY MEASURES?

Multivariable regression using reassessment star rating as the outcome, controlling for facility type, ownership, location and expectation

<table>
<thead>
<tr>
<th></th>
<th>Beta</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic infrastructure and equipment</td>
<td>0.98</td>
<td>(0.28, 1.68)</td>
</tr>
<tr>
<td>Management index</td>
<td>0.48</td>
<td>(0.07, 0.90)</td>
</tr>
<tr>
<td>Diagnosis and treatment accuracy</td>
<td>0.27</td>
<td>(-0.32, 0.86)</td>
</tr>
<tr>
<td>Good medical practice</td>
<td>0.89</td>
<td>(0.15, 1.62)</td>
</tr>
<tr>
<td>User experience</td>
<td>-0.17</td>
<td>(-0.93, 0.60)</td>
</tr>
</tbody>
</table>

- Basic infrastructure and equipment, management and good medical practice are associated with the star rating
- Good medical practice is not directly measured by Star Rating, but the totality of measures included in Star Rating are associated with good medical practice
- The management index was poorly correlated with domain A on management (.18)
- Diagnostic and treatment accuracy and user experience are not captured by the Star Rating
HOW DOES THE STAR RATING MAP TO OTHER QUALITY MEASURES?

- Good medical practice
- User experience
- Basic infrastructure and equipment
- Management index
- Diagnostic and treatment accuracy

For people:
- Processes of care
  - Competent care and systems
    - Positive user experience
- Quality impacts
  - Better health
    - Confidence in system
    - Economic benefit

Learning and improvement:
- Population health needs and expectations
- Governance policy, insurance, and non-health sectors
- Platforms accessibility and organisation of care
- Workforce numbers, skill, and support
- Tools equipment, medicines, and data

Equitable - Resilient - Efficient
INFRASTRUCTURE DOES NOT EQUAL QUALITY

RESEARCH QUESTION 3: HOW DOES THE STAR RATING INSTRUMENT PERFORM AS A QUALITY MEASUREMENT TOOL?

This is not just a quality measurement tool!

Yes and No!
CONCLUSIONS
- This evaluation suggests that the Star Rating Assessment is supporting primary care facilities to improve quality in Tanzania
- Targeted revisions to the instrument could make it even better
- The assessment should be integrated into other quality programs
NEXT STEPS

- Revising and piloting new tool
- Assess tertiary hospitals
- Next assessment soon
- Digital (offline)
ACKNOWLEDGMENTS

- Jeremy Veillard
- Margaret E. Kruk and Lab
- Peter Abdul DeWan
- Data Collectors
- Adelina Herman
- Pamela Sao