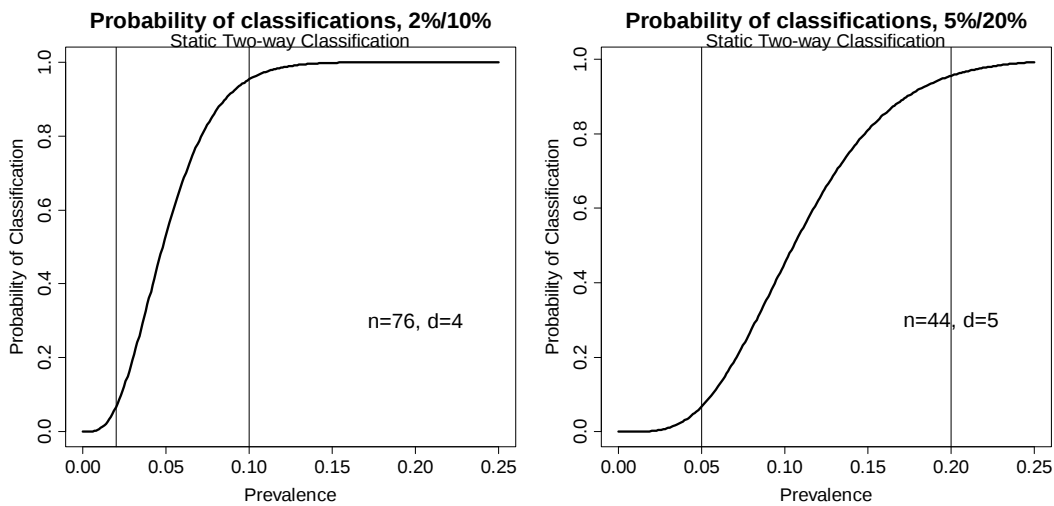


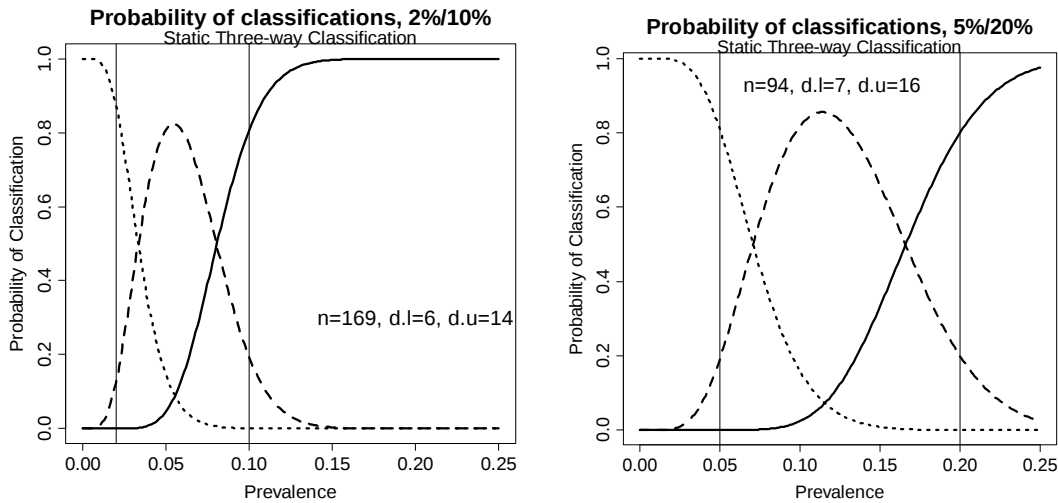
eAppendix

Operating characteristic curves for static two-way classification



eFigure1: The Operating Characteristic Curve for the static two-way classification schemes. The solid line is the probability of being classified as *high* DR prevalence at prevalences between 0-25% MDR TB.

Operating characteristic curves for static three-way classification



eFigure.2: The Operating Characteristic Curve for the static three-way classification schemes. The solid line is the probability of being classified as *high* MDR prevalence, long-dashed the probability as being classified as *moderate* MDR TB prevalence and the short dashed the probability as being classified as *low* MDR TB prevalence at prevalences between 0-25% MDR TB.

Development of the Truncated Sequential Sampling Design for Three-way Classification

When developing the design for the truncated sequential sampling plan for the three-way classification, we followed an approach similar to that outlined by Myatt et al (2008). One primary difference is that we did not assume any underlying distribution for DRTB prevalence, because there is not substantial *a priori* information on MDR prevalence. We simply evaluated the performance of the TSS design at the same lower, moderate, and upper thresholds as we used for the static three-way classification (0.02/0.06/0.10 or 0.05/0.125/0.20).

For each prevalence level and a given maximum sample size, we simulated a population with this underlying prevalence and randomly sampled the maximum sample size from this population. Starting at the minimum sample size, we counted the number of samples that are MDR infected. We compared this number to the lines generated for a fixed Alpha and Beta by the equations below.

$$\begin{aligned}d_{hn} &= \log(q_l/q_h)/\log((p_h*q_l)/(p_l*q_h))^n + \log(1-\text{Beta})/\text{Alpha}/\log((p_h*q_l)/(p_l*q_h)), \\ \text{and} \\ d_{ln} &= -\log(q_l/q_h)/\log((p_h*q_l)/(p_l*q_h))^n + \log(1-\text{Alpha})/\text{Beta}/\log((p_h*q_l)/(p_l*q_h))\end{aligned}$$

If the number of samples with MDR TB exceeded the upper line, then the site was classified as high; if the number was below the lower line, then the site was classified as low; otherwise, the site was not classified, and the next sample was considered. This process continued until reaching the maximum sample size. If the site remained unclassified at the maximum sample size, then it was classified in the moderate category. At each prevalence level, this process was repeated 10,000 times.

After preliminary explorations, we identified a set of Alphas, Betas, minimum sample sizes and maximum sample sizes that could plausibly meet our constraints. For each set of thresholds, we explored the same ranges of Alpha and Beta equal to 0.0001, 0.005 and 0.001. For the 2%/10% thresholds, we considered minimum sample sizes between 61 and 130 and maximum sample sizes between 151 and 190. For the 5%/20% thresholds, we considered minimum sample sizes between 31 and 60, and maximum sample sizes between 91 and 130.

We considered any design for which we correctly classified at the low DRTB threshold as low at least 80% of the time ($p_{ll} \geq 0.8$), the moderate threshold as moderate ($p_{mm} \geq 0.8$) and the high threshold as high at least 80% of the time ($p_{hh} \geq 0.8$). We also only considered designs that did not grossly misclassify sites at the low threshold as high more than 10% of the time ($p_{lh} \leq 0.1$), or sites at the high threshold as low more than 5% of the time ($p_{hl} \leq 0.05$). Amongst any eligible designs, we selected the design with the smallest maximum sample size. If multiple eligible sites had the smallest maximum sample size, then we chose the sites with the smallest minimum sample size.

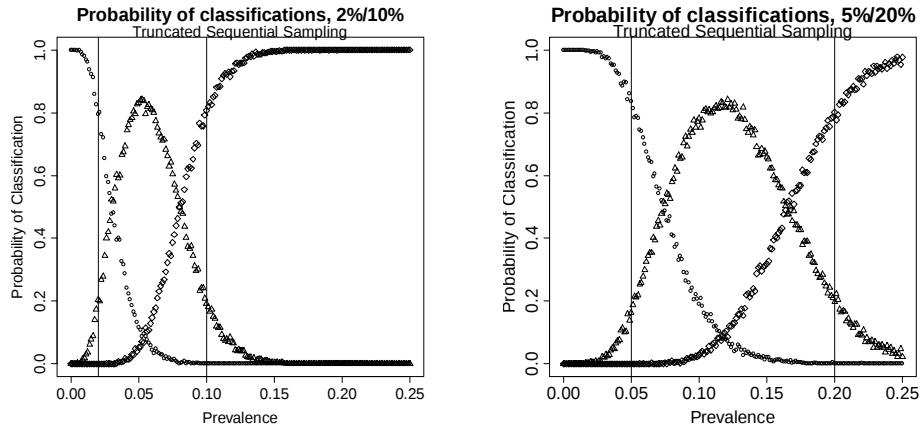
For the 2/10% thresholds, the following design was selected: minimum sample size of 62, a maximum sample size of 161, and $\text{Alpha}=0.0001$ and $\text{Beta} = 0.001$. For the 5/20% thresholds, the resulting design was a minimum sample size of 38, maximum sample size of 95, $\text{Alpha}=0.0001$ and $\text{Beta} = 0.001$. The tables to support the implementation of these designs are presented in Table A1 and A2.

We note that the Alpha and Betas have different meanings here than described previously in this paper, but we retain this notation because it reflects earlier descriptions of these methods. The definition of Alpha and Beta here reflect the classification errors of standard sequential sampling methods that result in two-way classifications.¹ Under standard sequential sampling methods, it is possible that sampling continues indefinitely (and may indeed never end) if numbers do not exceed the upper thresholds or fall below the lower threshold. In contrast, the truncated sequential design ensures that classification occurs within a fixed sample size. In the truncated sequential design, different values of Alpha and Beta are explored for the threshold lines; the recommended values of Alpha and Beta are determined as those that result in a classification system which meets desired operational characteristics. In our case, where we examine three-way classification for the truncated sequential sampling, the desired characteristics are: 1) at least 80% probability of correct classification ($p_{ll}=p_{mm}=p_{hh}=0.80$) at each of the three levels (low, moderate and high), 2) no more than 10% probability of gross misclassification of low prevalence as high ($p_{lh}=0.1$), and 3) no more than 5% probability of gross misclassification of high prevalence as low ($p_{hl}=0.05$). These performance metrics are the same as the ones we set for our static three-way classification and are comparable to our two-way classification approach. As with the other systems, there are multiple options that meet these constraints, so we select the one with the smallest average sample size.

Operating Characteristic Curves of the Truncated Sequential Sampling Design for Three-way Classification

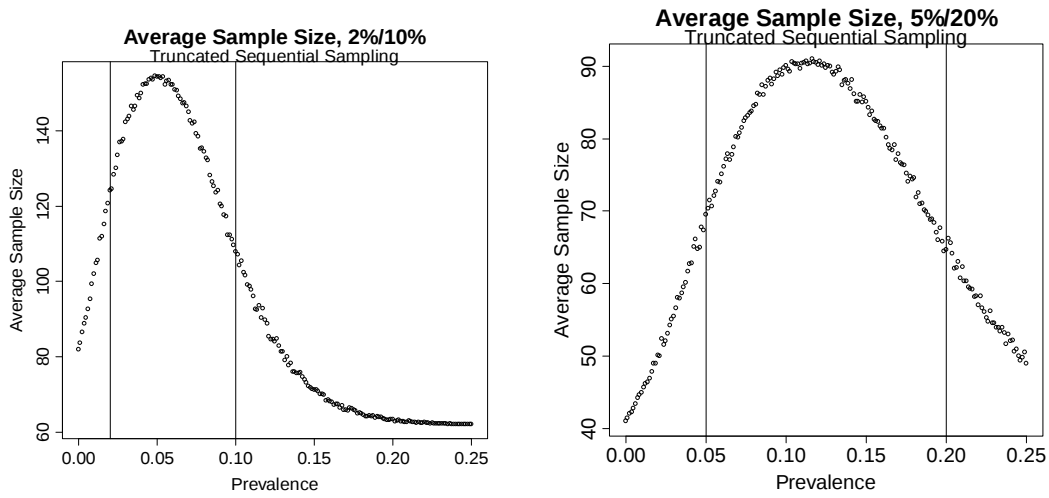
Following the TSS design recommended from the above simulations, we classified the sample as *low*, *moderate*, or *high* MDR TB prevalence. The OCC below result from the 1000 simulations repeated at each prevalence level (for prevalences between 0-25%), where we calculated the probability of being classified at each level.

¹ Wald A. 1947. Sequential Analysis. New York: John Wiley and Sons.



eFigure 3: The Operating Characteristic Curve for the TSS three-way classification schemes. The diamonds are the probability of being classified as *high* MDR prevalence, the triangles are the probability as being classified as *moderate* MDR TB prevalence and the dots dashed the probability as being classified as *low* MDR TB prevalence at prevalences between 0-25% MDR TB.

For each of the above simulations, we also tracked the number of samples required to reach at the appropriate classifications. The plots below show the average sample sizes as a function of prevalence.



eFigure.4: The average sample size by prevalence of the TSS three-way classifications.

Sample Size	Lower Cutoff	Number Observed	Upper Cutoff		Sample Size	Lower Cutoff	Number Observed	Upper Cutoff
62	NA	_____	8		112	2	_____	11
63	NA	_____	8		113	2	_____	11
64	NA	_____	8		114	2	_____	11
65	NA	_____	8		115	2	_____	11
66	NA	_____	8		116	2	_____	11
67	NA	_____	8		117	2	_____	11
68	NA	_____	8		118	2	_____	11
69	NA	_____	8		119	2	_____	11
70	NA	_____	8		120	2	_____	11
71	NA	_____	9		121	3	_____	11
72	NA	_____	9		122	3	_____	11
73	NA	_____	9		123	3	_____	11
74	NA	_____	9		124	3	_____	11
75	NA	_____	9		125	3	_____	11
76	NA	_____	9		126	3	_____	11
77	NA	_____	9		127	3	_____	11
78	NA	_____	9		128	3	_____	11
79	NA	_____	9		129	3	_____	11
80	NA	_____	9		130	3	_____	11
81	NA	_____	9		131	3	_____	12
82	1	_____	9		132	3	_____	12
83	1	_____	9		133	3	_____	12
84	1	_____	9		134	3	_____	12
85	1	_____	9		135	3	_____	12
86	1	_____	9		136	3	_____	12
87	1	_____	9		137	3	_____	12
88	1	_____	9		138	3	_____	12
89	1	_____	9		139	3	_____	12
90	1	_____	9		140	3	_____	12
91	1	_____	10		141	4	_____	12
92	1	_____	10		142	4	_____	12
93	1	_____	10		143	4	_____	12
94	1	_____	10		144	4	_____	12
95	1	_____	10		145	4	_____	12
96	1	_____	10		146	4	_____	12
97	1	_____	10		147	4	_____	12
98	1	_____	10		148	4	_____	12
99	1	_____	10		149	4	_____	12
100	1	_____	10		150	4	_____	12
101	1	_____	10		151	4	_____	13
102	2	_____	10		152	4	_____	13
103	2	_____	10		153	4	_____	13
104	2	_____	10		154	4	_____	13
105	2	_____	10		155	4	_____	13
106	2	_____	10		156	4	_____	13
107	2	_____	10		157	4	_____	13
108	2	_____	10		158	4	_____	13
109	2	_____	10		159	4	_____	13
110	2	_____	10		160	4	_____	13
111	2	_____	11		161	5	_____	13

eTable 1: Truncated Sequential Sample table for the 2%/10% thresholds

Sample Size	Lower Cutoff	Number Observed	Upper Cutoff		Sample Size	Lower Cutoff	Number Observed	Upper Cutoff
38	NA	_____	10		67	3	_____	13
39	NA	_____	10		68	4	_____	13
40	NA	_____	10		69	4	_____	13
41	1	_____	10		70	4	_____	13
42	1	_____	10		71	4	_____	13
43	1	_____	10		72	4	_____	13
44	1	_____	10		73	4	_____	13
45	1	_____	10		74	4	_____	14
46	1	_____	10		75	4	_____	14
47	1	_____	11		76	4	_____	14
48	1	_____	11		77	5	_____	14
49	1	_____	11		78	5	_____	14
50	2	_____	11		79	5	_____	14
51	2	_____	11		80	5	_____	14
52	2	_____	11		81	5	_____	14
53	2	_____	11		82	5	_____	14
54	2	_____	11		83	5	_____	15
55	2	_____	11		84	5	_____	15
56	2	_____	12		85	5	_____	15
57	2	_____	12		86	6	_____	15
58	2	_____	12		87	6	_____	15
59	3	_____	12		88	6	_____	15
60	3	_____	12		89	6	_____	15
61	3	_____	12		90	6	_____	15
62	3	_____	12		91	6	_____	15
63	3	_____	12		92	6	_____	16
64	3	_____	12		93	6	_____	16
65	3	_____	13		94	6	_____	16
66	3	_____	13		95	7	_____	16

eTable 2: Truncated Sequential Sample table for the 5%/20% thresholds

[illegible]

35	NA	NA	NA	NA	NA	NA	NA	NA	NA
36	NA	NA	NA	NA	NA	NA	NA	NA	NA
37	NA	NA	NA	NA	NA	NA	NA	NA	NA
38	NA	NA	NA	NA	NA	NA	NA	NA	NA
41	NA	NA	NA	NA	NA	NA	NA	NA	NA
43	NA	NA	NA	NA	NA	NA	NA	NA	NA
44	NA	NA	NA	NA	NA	NA	NA	NA	NA
45	NA	NA	NA	NA	NA	NA	NA	NA	NA
47	NA	NA	NA	NA	NA	NA	NA	NA	NA
48	NA	NA	NA	NA	NA	NA	NA	NA	NA
49	NA	NA	NA	NA	NA	NA	NA	NA	NA
50	NA	NA	NA	NA	NA	NA	NA	NA	NA
51	NA	NA	NA	NA	NA	NA	NA	NA	NA
52	NA	NA	NA	NA	NA	NA	NA	NA	NA
53	NA	NA	NA	NA	NA	NA	NA	NA	NA
54	NA	NA	NA	NA	NA	NA	NA	NA	NA
55	NA	NA	NA	NA	NA	NA	NA	NA	NA
56	NA	NA	NA	NA	NA	NA	NA	NA	NA
57	NA	NA	NA	NA	NA	NA	NA	NA	NA
58	8	4	2	23	14	2	NA	NA	NA
59	NA	NA	NA	NA	NA	NA	NA	NA	NA
60	NA	NA	NA	NA	NA	NA	NA	NA	NA

NA = classification not possible before exhausting the samples available

eTable 3: LQAS classifications for Ukraine centers at the 2%/10% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Group Number	Sample Size (max=76)	Number with resistance	Classification	Sample Size (max=169)	Number with resistance	Classification	Sample Size (max=161)	Number with resistance	Classification
1	30	4	2	116	14	2	83	10	2
2	26	4	2	58	14	2	62	14	2
3	47	4	2	111	14	2	81	10	2
4	10	4	2	47	14	2	62	23	2
5	39	4	2	108	14	2	62	10	2

NA = classification not possible before exhausting the samples available

eTable 4: LQAS classifications for Ukraine groups at the 2%/10% thresholds.

[illegible]

35	NA	NA	NA	NA	NA	NA	NA	NA	NA
36	NA	NA	NA	NA	NA	NA	NA	NA	NA
37	NA	NA	NA	NA	NA	NA	NA	NA	NA
38	NA	NA	NA	NA	NA	NA	NA	NA	NA
41	NA	NA	NA	NA	NA	NA	NA	NA	NA
43	NA	NA	NA	NA	NA	NA	NA	NA	NA
44	NA	NA	NA	NA	NA	NA	NA	NA	NA
45	NA	NA	NA	NA	NA	NA	NA	NA	NA
47	NA	NA	NA	NA	NA	NA	NA	NA	NA
48	NA	NA	NA	NA	NA	NA	NA	NA	NA
49	NA	NA	NA	NA	NA	NA	NA	NA	NA
50	NA	NA	NA	NA	NA	NA	NA	NA	NA
51	NA	NA	NA	NA	NA	NA	NA	NA	NA
52	NA	NA	NA	NA	NA	NA	NA	NA	NA
53	NA	NA	NA	NA	NA	NA	NA	NA	NA
54	NA	NA	NA	NA	NA	NA	NA	NA	NA
55	NA	NA	NA	NA	NA	NA	NA	NA	NA
56	NA	NA	NA	NA	NA	NA	NA	NA	NA
57	NA	NA	NA	NA	NA	NA	NA	NA	NA
58	11	5	2	27	16	2	NA	NA	NA
59	NA	NA	NA	NA	NA	NA	NA	NA	NA
60	NA	NA	NA	NA	NA	NA	NA	NA	NA

NA = classification not possible before exhausting the samples available

eTable 5: LQAS classifications for Ukraine centers at the 5%/20% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Group Number	Sample Size (max=44)	Number with resistance	Classification	Sample Size (max=94)	Number with resistance	Classification	Sample Size (max=95)	Number with resistance	Classification
1	36	5	2	94	12	1	95	12	1
2	27	5	2	77	16	2	43	11	2
3	43	3	0	94	11	1	95	11	1
4	12	5	2	49	16	2	38	12	2
5	40	5	2	94	11	1	95	12	1

NA = classification not possible before exhausting the samples available

eTable 6: LQAS classifications for Ukraine groups at the 5%/20% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Center Number	Sample Size (max=76)	Number with resistance	Classification	Sample Size (max=169)	Number with resistance	Classification	Sample Size (max=161)	Number with resistance	Classification
1	NA	NA	NA	NA	NA	NA	NA	NA	NA
2	NA	NA	NA	NA	NA	NA	NA	NA	NA
3	NA	NA	NA	NA	NA	NA	NA	NA	NA
4	NA	NA	NA	NA	NA	NA	NA	NA	NA
5	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	NA	NA	NA	NA	NA	NA	NA	NA	NA
7	NA	NA	NA	NA	NA	NA	NA	NA	NA
8	NA	NA	NA	NA	NA	NA	NA	NA	NA
9	73	0	0	NA	NA	NA	82	0	0
10	74	1	0	NA	NA	NA	NA	NA	NA
11	NA	NA	NA	NA	NA	NA	NA	NA	NA
12	NA	NA	NA	NA	NA	NA	NA	NA	NA
13	73	0	0	NA	NA	NA	NA	NA	NA
14	NA	NA	NA	NA	NA	NA	NA	NA	NA
15	NA	NA	NA	NA	NA	NA	NA	NA	NA
16	NA	NA	NA	NA	NA	NA	NA	NA	NA
17	NA	NA	NA	NA	NA	NA	NA	NA	NA
18	75	2	0	NA	NA	NA	NA	NA	NA
19	NA	NA	NA	NA	NA	NA	NA	NA	NA
20	NA	NA	NA	NA	NA	NA	NA	NA	NA
21	NA	NA	NA	NA	NA	NA	NA	NA	NA
22	NA	NA	NA	NA	NA	NA	NA	NA	NA
23	NA	NA	NA	NA	NA	NA	NA	NA	NA
24	NA	NA	NA	NA	NA	NA	NA	NA	NA

NA = classification not possible before exhausting the samples available

eTable 7: LQAS classifications for Tanzania centers at the 2%/10% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Group Number	Sample Size (max=76)	Number with resistance	Classification	Sample Size (max=169)	Number with resistance	Classification	Sample Size (max=161)	Number with resistance	Classification
1	73	0	0	NA	NA	NA	82	0	0
2	75	2	0	166	2	0	121	2	0
3	74	1	0	167	3	0	121	2	0
4	74	1	0	165	1	0	102	1	0
5	73	0	0	NA	NA	NA	NA	NA	NA

NA = classification not possible before exhausting the samples available

eTable 8: LQAS classifications for Tanzania groups at the 2%/10% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Center Number	Sample Size (max=44)	Number with resistance	Classification	Sample Size (max=94)	Number with resistance	Classification	Sample Size (max=95)	Number with resistance	Classification
1	NA	NA	NA	NA	NA	NA	NA	NA	NA
2	NA	NA	NA	NA	NA	NA	NA	NA	NA
3	40	0	0	NA	NA	NA	41	0	0
4	NA	NA	NA	NA	NA	NA	NA	NA	NA
5	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	41	1	0	NA	NA	NA	50	1	0
7	40	0	0	NA	NA	NA	50	1	0
8	40	0	0	NA	NA	NA	41	0	0
9	40	0	0	NA	NA	NA	41	0	0
10	41	1	0	NA	NA	NA	50	1	0
11	NA	NA	NA	NA	NA	NA	NA	NA	NA
12	NA	NA	NA	NA	NA	NA	NA	NA	NA
13	40	0	0	NA	NA	NA	41	0	0
14	41	1	0	NA	NA	NA	50	1	0
15	40	0	0	NA	NA	NA	NA	NA	NA
16	NA	NA	NA	NA	NA	NA	NA	NA	NA
17	NA	NA	NA	NA	NA	NA	NA	NA	NA
18	41	1	0	NA	NA	NA	59	2	0
19	NA	NA	NA	NA	NA	NA	NA	NA	NA
20	NA	NA	NA	NA	NA	NA	NA	NA	NA
21	NA	NA	NA	NA	NA	NA	NA	NA	NA
22	NA	NA	NA	NA	NA	NA	NA	NA	NA
23	NA	NA	NA	NA	NA	NA	NA	NA	NA
24	NA	NA	NA	NA	NA	NA	NA	NA	NA

NA = classification not possible before exhausting the samples available

eTable 9: LQAS classifications for Tanzania centers at the 5%/20% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Group Number	Sample Size (max=44)	Number with resistance	Classification	Sample Size (max=94)	Number with resistance	Classification	Sample Size (max=95)	Number with resistance	Classification
1	40	0	0	88	0	0	41	0	0
2	41	1	0	90	2	0	59	2	0
3	41	1	0	89	1	0	50	1	0
4	40	0	0	89	1	0	41	0	0
5	40	0	0	NA	NA	NA	41	0	0

NA = classification not possible before exhausting the samples available

eTable 10: LQAS classifications for Tanzania groups at the 5%/20% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Center Number	Sample Size (max=76)	Number with resistance	Classification	Sample Size (max=169)	Number with resistance	Classification	Sample Size (max=161)	Number with resistance	Classification
25	76	3	0	NA	NA	NA	NA	NA	NA
26	NA	NA	NA	NA	NA	NA	NA	NA	NA
27	NA	NA	NA	NA	NA	NA	NA	NA	NA
28	NA	NA	NA	NA	NA	NA	NA	NA	NA
29	NA	NA	NA	NA	NA	NA	NA	NA	NA
30	NA	NA	NA	NA	NA	NA	NA	NA	NA
31	NA	NA	NA	NA	NA	NA	NA	NA	NA
32	NA	NA	NA	NA	NA	NA	NA	NA	NA
33	74	1	0	NA	NA	NA	NA	NA	NA
34	NA	NA	NA	NA	NA	NA	NA	NA	NA
35	NA	NA	NA	NA	NA	NA	NA	NA	NA
36	NA	NA	NA	NA	NA	NA	NA	NA	NA
37	NA	NA	NA	NA	NA	NA	NA	NA	NA
38	NA	NA	NA	NA	NA	NA	NA	NA	NA
39	25	4	2	NA	NA	NA	NA	NA	NA
40	NA	NA	NA	NA	NA	NA	NA	NA	NA
41	NA	NA	NA	NA	NA	NA	NA	NA	NA
42	NA	NA	NA	NA	NA	NA	NA	NA	NA
43	NA	NA	NA	NA	NA	NA	NA	NA	NA
44	75	2	0	168	4	0	121	2	0

NA = classification not possible before exhausting the samples available

eTable11: LQAS classifications for Vietnam centers at the 2%/10% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Group Number	Sample Size (max=76)	Number with resistance	Classification	Sample Size (max=169)	Number with resistance	Classification	Sample Size (max=161)	Number with resistance	Classification
1	75	2	0	166	2	0	121	2	0
2	73	0	0	NA	NA	NA	82	0	0
3	56	4	2	NA	NA	NA	NA	NA	NA
4	NA	NA	NA	NA	NA	NA	NA	NA	NA
5	75	2	0	168	4	0	121	2	0

NA = classification not possible before exhausting the samples available

eTable 12: LQAS classifications for Vietnam groups at the 2%/10% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Center Number	Sample Size (max=44)	Number with resistance	Classification	Sample Size (max=94)	Number with resistance	Classification	Sample Size (max=95)	Number with resistance	Classification
25	41	1	0	94	4	1	68	3	0
26	NA	NA	NA	NA	NA	NA	NA	NA	NA
27	NA	NA	NA	NA	NA	NA	NA	NA	NA
28	NA	NA	NA	NA	NA	NA	NA	NA	NA
29	NA	NA	NA	NA	NA	NA	NA	NA	NA
30	NA	NA	NA	NA	NA	NA	NA	NA	NA
31	41	1	0	NA	NA	NA	59	2	0
32	40	0	0	NA	NA	NA	41	0	0
33	40	0	0	89	1	0	41	0	0
34	40	0	0	NA	NA	NA	41	0	0
35	41	1	0	NA	NA	NA	NA	NA	NA
36	NA	NA	NA	NA	NA	NA	NA	NA	NA
37	NA	NA	NA	NA	NA	NA	NA	NA	NA
38	NA	NA	NA	NA	NA	NA	NA	NA	NA
39	32	5	2	NA	NA	NA	NA	NA	NA
40	40	0	0	NA	NA	NA	41	0	0
41	NA	NA	NA	NA	NA	NA	NA	NA	NA
42	NA	NA	NA	NA	NA	NA	NA	NA	NA
43	NA	NA	NA	NA	NA	NA	NA	NA	NA
44	41	1	0	90	2	0	59	2	0

NA = classification not possible before exhausting the samples available

eTable 13: LQAS classifications for Vietnam centers at the 5%/20% thresholds.

	Two-way classification			Three-way classification			TSS Three-way classification		
Group Number	Sample Size (max=44)	Number with resistance	Classification	Sample Size (max=94)	Number with resistance	Classification	Sample Size (max=95)	Number with resistance	Classification
1	41	1	0	90	2	0	59	2	0
2	40	0	0	88	0	0	41	0	0
3	43	3	0	94	6	0	77	4	0
4	40	0	0	NA	NA	NA	41	0	0
5	41	1	0	90	2	0	59	2	0

NA = classification not possible before exhausting the samples available

eTable 14: LQAS classifications for Vietnam groups at the 5%/20% thresholds.