Who is Making the Grade?

Statistical Methods for Detecting Unusual Performance in HIV Care and Treatment Programs Using EMR Data in a Low-Resource Setting

Adrien Allorant – Department of Global Health (DGH) & IHME
University of Washington (UW)
allorant@uw.edu
Nancy Puttkammer – DGH & I-TECH (UW)

CENTER FOR STUDIES IN DEMOGRAPHY & ECOLOGY

Background

Public and private sponsors investing in improving health in low and middle income countries (LMICs) often use defined quantitative indicators and targets to gauge program progress at the health-facility level. However, these methods fail to address reasons which could explain variability in performance over time, such as a varying risk profile of patients, sampling variability, and variable data quality. Statistical methods accounting for these factors exist nevertheless and are routinely used, for instance in the UK[1,2]. We used *iSanté*, Haiti's Electronic Medical Record (EMR) system, to demonstrate the pertinence of using these statistical methods on process and outcome indicators of HIV care and treatment guidelines to detect unusual performance in quality of care (QoC) among health facilities, in a low-resource setting.

Objectives

1. To construct indicators of QoC on 3 process and 3 outcome indicators (described in Table 1)
2. To construct a composite measure of QoC combining the evidence of each indicator

Methods

Study design

This secondary analysis involved data from 65,472 patients seeking HIV/AIDS care and treatment services in 84 health facilities from June 2016 to March 2018. We used case-mix adjustment including demographic and baseline clinical variables to adjust our outcome indicators for patients' characteristics.

	Indicators of QoC	Definition
	Process indicators	
M1	Uptake of 'Test and Start' approach	Patient started Antiretroviral treatment (ART) within 1 month of diagnosis
M2	Uptake of viral load testing	Patient up-to-date with respect to viral load testing
M3	Uptake of multi-month scripting (MMS)	ART prescription intervals of greater than 35 days per guidelines, based on definition of stable patients who are eligible for MMS
	Outcome indicators	
M4	ART 30-day retention	Patient has returned within 30 days of expected ART refill date
M5	ART 6-month retention non-pregnant adults	Patient has returned within 30 days of expected ART refill date, 6 months after starting ART
M6	ART 6-month retention pregnant women	Patient has returned within 30 days of expected ART refill date, 6 months after starting ART AND patient was pregnant or post-partum at time of starting ART

Table 1: Definition of the indicators of Quality of Care

Patient-level model

We modelled patient *j*'s retention on ART - indicators 4 to 6 (**M4-M6**) - using a logistic regression:

$$Retention_j \sim \text{Bernouilli}\left(logit^{-1}(\boldsymbol{x}_j\beta)\right)$$

where $\mathbf{x}_j$ is a vector of covariates including age, gender, marital status, HIV status of partner, date of first diagnosis, age at ART start, years on ART, and average travel times to the facility.
We calculated the predicted value of retention of each patient *j*, and aggregated these probabilities over each facility *i*, to obtain E_i the expected number of patients retained in facility *i*.

Facility-level model

For **M4-M6**, we assumed that the observed number of patients retained on ART in facility *i*, O_i, followed a Poisson distribution of parameter E_i, a usual assumption for a standardized ratio:

$$O_i \sim \text{Poisson}\left(E_i\right)$$

For **M1-M3**, we recorded, in each facility *i* with N_i patients, the number of times the corresponding standard care had been respected Y_i. We expected some variabilities across facilities. However, we expected the proportions of *success* to fluctuate around a common national average:

$$Y_i \sim \text{Binomial}\left(\pi_0, N_i\right)$$

Z-scores represent the deviation from a standard on a common scale; transforming the indicator and the standard before conversion to a Z-score is common. Following Spiegelhalter *et al.*[2], we used an inverse sine and a square-root transformation, for our process (**M1-M3**), and outcome indicators (**M4-M6**), respectively:

$$z = 2\sqrt{n}\left(\sin^{-1}\left(\sqrt{p}\right) - \sin^{-1}\left(\sqrt{\theta_0}\right)\right) \qquad \text{(M1-M3)}$$

$$z = 2\left(\sqrt{O} - \sqrt{E}\right) \qquad \text{(M4-M6)}$$

Composite Z-scores combine the evidence from each of the 6 indicators to detect consistent patterns of unusual performance. We followed Bardsley *et al.*[1] to calculate for each facility *i* a composite Z-score:

$$Z_i^{composite} = \frac{\sum_{k=1}^{k=6} w_k z_{ik}}{\sqrt{\sum_k \sum_j w_k w_j c_{kj}}}$$

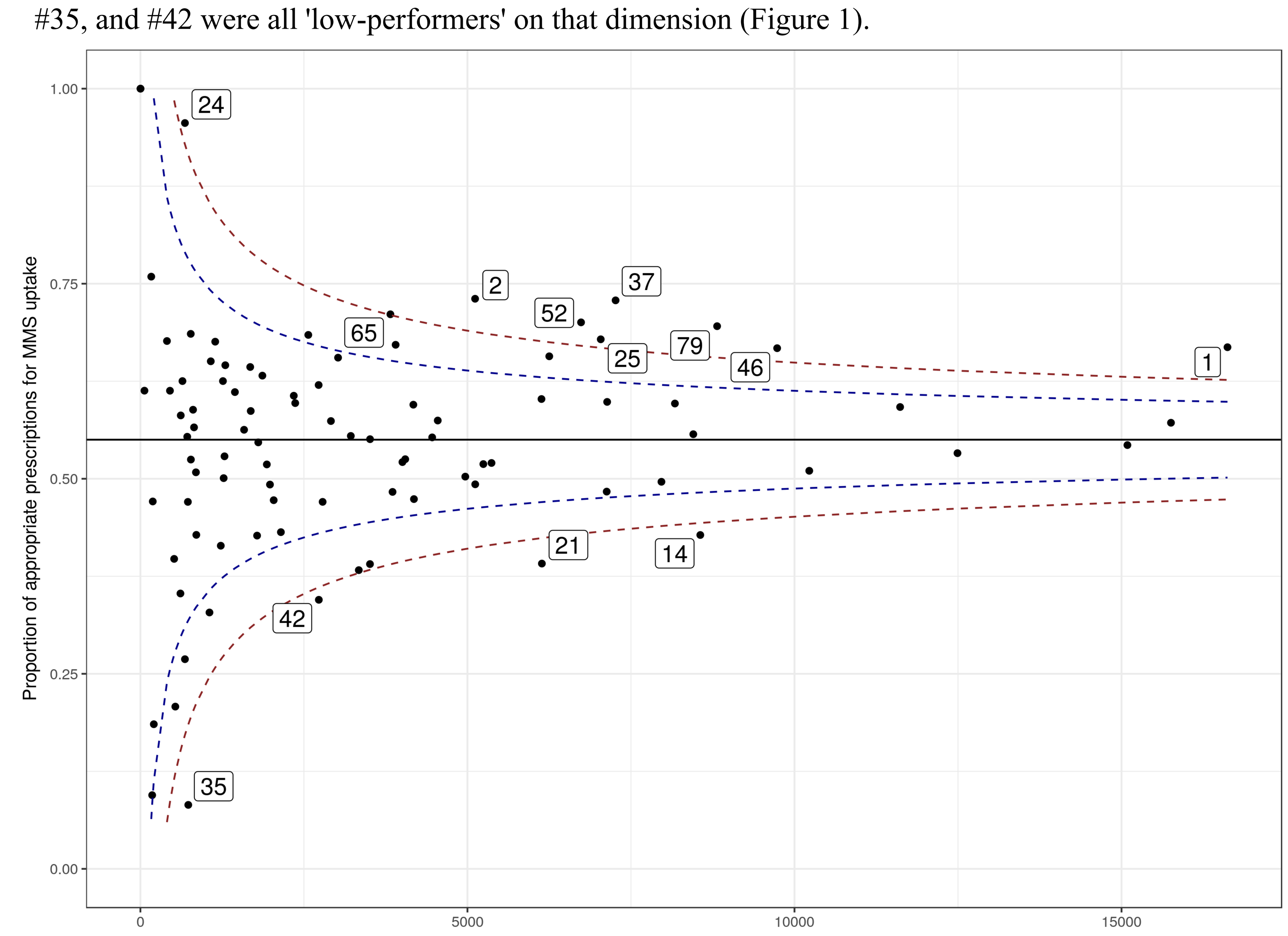
where c_{kj} are the correlations between the z-scores, and $c_{kk} = 1$

The current suggestion is to use the following weights: $w_k = 1/\sum_j c_{kj}$, to down weight pairs of highly correlated z-scores[2].

Overdispersion occurs when the within-facility variability is underestimated and leads to an inappropriately high number of outliers[2]. We can estimate an overdispersion factor: $\widehat{\phi} = \sum_{i=1}^{I} z_i^2 / I$, where *I* represents the number of facilities, and z_i the z-score of facility *i*.
This factor is used to inflate the null standard error in order to avoid underestimating the within-facility sampling error: $\widetilde{s_0} = s_0 \times \sqrt{\widehat{\phi}}$

Results

Funnel plots are a convenient tool to detect unusual performance[2,3]. For instance, facilities #1, #2, #24, #25, #37, #46, #52, and #79 all displayed a proportion of appropriate use of MMS significantly higher than the national average, and can be considered 'star performers' falling above the 99.8% confidence limit of expected performance on that dimension of QoC. Facilities #14, #21, #35, and #42 were all 'low-performers' on that dimension (Figure 1).



We identified several star and poor performers for each QoC indicator using funnel plots. However, our composite Z-score identified 3 facilities with consistent evidence of unusually poor performance - #81, #24, and #35, and 2 'star performers' - #1, and #37) (Table 2).

	M1	M2	M3	M4	M5	M6	Composite_Zscore
1	0.33	-2.40	-2.39	-2.07	-1.51	-1.92	-2.27
37	0.08	-1.58	-2.43	-1.78	-1.09	-1.51	-2.05
79	-0.00	-1.49	-2.14	-1.61	-1.30	-1.73	-1.95
2	-1.61	-0.66	-2.05	-0.20	0.51	0.54	-1.88
61	-1.63	-0.52	-1.02	-0.59	-0.20	-0.03	-1.69
26	-1.14	-1.28	-0.56	-0.55	-0.51	-0.55	-1.49
46	0.35	-2.04	-1.79	-1.09	0.41	-0.76	-1.36
36	-0.27	-1.30	-0.37	-1.15	-1.56	-1.13	-1.23
52	0.22	1.25	-1.93	-2.22	-1.71	-1.51	-1.19
65	-1.08	-0.71	-1.54	0.79	0.06	-0.21	-1.19
70	0.70	0.30	1.42	0.24	1.42	0.26	1.28
5	0.79	0.38	-0.03	0.70	4.05	3.67	1.33
56	2.22	0.19	0.02	-0.05	-0.00	0.15	1.41
18	1.14	-0.37	0.75	0.82	2.78	0.85	1.45
17	2.08	0.39	0.08	0.84	-0.42	-0.15	1.60
42	1.79	-0.22	1.82	0.46	-0.11	1.59	1.98
20	3.71	0.23	0.66	-0.76	0.48	0.41	1.99
81	2.98	-0.01	0.99	-0.32	-0.76	0.33	2.04
24	3.09	0.27	-2.07	4.24	1.35	0.30	2.16
35	1.68	0.95	2.43	1.48	2.09	2.31	3.03

Conclusion

Our work demonstrates the potential of routine data systems such as *iSanté* for institutional performance monitoring. The rich person-level data and robust statistical methods allow detection of star and weak performers, among all health facilities. These results could be leveraged for: a) Risk-based targeted inspection of facilities; b) Evidence-based health systems strengthening and funding allocation decisions.

References
[1] Bardsley *et al.*, *Qual Saf Health Care*, **2009**,18(3) ; [2] Spiegelhalter *et al.*, *J R Soc Stat Ser A Stat Soc*, **2012**,175(1); [3] Spiegelhalter *et al.*, Stat Med, **2005**,24(8)

Acknowledgements
This research was undertaken in collaboration with the Ministère de la Santé Publique et de la Population (MSPP) of Haiti and the US Centers for Disease Control and Prevention (CDC). The work has been supported by the President's Emergency Plan for AIDS Relief (PEPFAR) through the US Centers for Disease Control and Prevention (https://www.cdc.gov/), under award number NU2GGH001130-04-00, to the International Training and Education Center for Health (I-TECH) at the University of Washington and by NIAID, NCI, NIMH, NIDA, NICHD, NHLBI, NIA, NIGMS, NIDDK of the National Institutes of Health (https://www.nih.gov/) under award number AI027757 to the University of Washington Center for AIDS Research (CFAR). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the National Institutes of Health.