HIV Quality Report Cards: Impact of Case-Mix Adjustment and Statistical Methods

Michael E. Ohl,[1,2] Kelly K. Richardson,[1,2] Michihiko Goto,[1,2] Mary Vaughan-Sarrazin,[1,2] Marin L. Schweizer,[1,2] and Eli N. Perencevich[1,2]

[1]Center for Comprehensive Access and Delivery Research and Evaluation, Iowa City Veterans Affairs Medical Center, and [2]Department of Internal Medicine, University of Iowa Carver College of Medicine, Iowa City

Background. There will be increasing pressure to publicly report and rank the performance of healthcare systems on human immunodeficiency virus (HIV) quality measures. To inform discussion of public reporting, we evaluated the influence of case-mix adjustment when ranking individual care systems on the viral control quality measure.

Methods. We used data from the Veterans Health Administration (VHA) HIV Clinical Case Registry and administrative databases to estimate case-mix adjusted viral control for 91 local systems caring for 12 368 patients. We compared results using 2 adjustment methods, the observed-to-expected estimator and the risk-standardized ratio.

Results. Overall, 10 913 patients (88.2%) achieved viral control (viral load ≤400 copies/mL). Prior to case-mix adjustment, system-level viral control ranged from 51% to 100%. Seventeen (19%) systems were labeled as low outliers (performance significantly below the overall mean) and 11 (12%) as high outliers. Adjustment for case mix (patient demographics, comorbidity, CD4 nadir, time on therapy, and income from VHA administrative databases) reduced the number of low outliers by approximately one-third, but results differed by method. The adjustment model had moderate discrimination (c statistic = 0.66), suggesting potential for unadjusted risk when using administrative data to measure case mix.

Conclusions. Case-mix adjustment affects rankings of care systems on the viral control quality measure. Given the sensitivity of rankings to selection of case-mix adjustment methods—and potential for unadjusted risk when using variables limited to current administrative databases—the HIV care community should explore optimal methods for case-mix adjustment before moving forward with public reporting.

Keywords. HIV; quality measurement; public reporting; electronic health records.

In an effort to standardize quality measurement in human immunodeficiency virus (HIV) care, the National Committee for Quality Assurance (NCQA) recently released HIV quality measures that include indices of viral control at the provider and care-system levels [1]. At the system level, viral control is essentially defined as the proportion of patients receiving potent antiretroviral therapy (ART) who have an HIV RNA below the limit of quantification.

There will likely be increasing pressure to publicly report and rank the performance of individual HIV care systems on the viral control quality measure. Publication of quality report cards is already common for other conditions, such as acute myocardial infarction, congestive heart failure, pneumonia, and various surgical procedures [2]. These report cards profile performance of individual healthcare systems on established quality measures, and identify systems that are outliers with performance significantly above or below the population average for the measure, or some other threshold that is deemed to represent acceptable performance. The intent is that the external accountability created by public reporting will incent care systems to pursue quality improvement, provide a foundation for value-based

Received 22 April 2014; accepted 2 July 2014; electronically published 16 July 2014.

Correspondence: Michael E. Ohl, MD, MSPH, Division of Infectious Diseases, University of Iowa Carver College of Medicine, Center for Comprehensive Access and Delivery Research and Evaluation (CADRE), MS 152, Iowa City VAMC, Iowa City, IA 52246 (michael-ohl@uiowa.edu).

Clinical Infectious Diseases 2014;59(8):1160–7

DOI: 10.1093/cid/ciu551

purchasing by payers (ie, pay for performance), and inform patients in choosing high-quality care sites.

If reporting of viral control for specific healthcare systems is to serve these purposes, then the reported measure should reflect the quality of care delivered by the system, and not simply the characteristics of patients in care (ie, case mix). If case mix is a major determinant of viral control, then systems that care for patients with higher rates of risk factors associated with poor control, such as mental health and substance use comorbidities or viral resistance, will be misrepresented in report cards [3, 4]. Currently, little is known about the extent to which case-mix adjustment modifies the viral control measure for care systems.

When deciding to adjust for case mix, it is necessary to make decisions on 2 critical and related methodological issues. First, one must select the variables that will represent patient characteristics in adjustment models. These should accurately capture patient characteristics that both influence viral control and vary in frequency across healthcare systems. Ideally, patient variables should be routinely collected in administrative databases available to reporting agencies. Second, one must choose which statistical method to apply when using patient variables to adjust for case mix. When adjusting quality measures for case mix, choice of statistical method influences both the estimated performance of individual care systems and which are labeled as outliers [5, 6]. This remains true even when identical patient variables are used. The relative merits of different methods for case-mix adjustment have been extensively discussed in the health services literature, but there is less awareness of these issues in the HIV care community, which has much at stake as public reporting moves forward [5–9].

In an effort to initiate and inform a discussion about public reporting in HIV care, we used data from the national Veterans Health Administration (VHA) HIV clinical case registry (CCR) and administrative databases to create viral control report cards [10]. We then examined the impacts of case-mix adjustment and choice of statistical method on reported performance for each system. Specifically, we compared 2 methods that have been commonly used in public reporting: (1) the observed over expected (O/E) estimator previously used by the New York Coronary Artery Bypass Graft and other public reporting programs, and (2) the risk-standardized ratio (RSR) currently used by the Centers for Medicare and Medicaid Services (CMS) in the Hospital Compare program [2, 11].

METHODS

Data Sources and Patient Population

CCR identifies patients in care for HIV infection and compiles electronic health record data from the Veterans Health Information System and Technology Architecture. Data elements include patient demographics, laboratory values, medications, and diagnoses by *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) codes [10]. Using 2009 data, we merged CCR files with VHA outpatient care files that included information on site and type of clinic visits and patient income.

We defined an HIV care system as a single VHA infectious disease specialty clinic and limited the sample to patients who met criteria for inclusion in the NCQA viral control measure denominator, specifically (1) receipt of potent ART for at least 6 months and (2) at least 2 clinic visits in 2009 at least 60 days apart [1]. Potent ART was defined as simultaneous receipt of least 3 medications from at least 2 classes, or coformulated abacavir/lamivudine/zidovudine (Trizivir). Ritonavir prescribed at doses used for pharmacologic boosting was not considered a separate medication. In line with protocols used by CMS for Hospital Compare report cards, we excluded care systems with <25 eligible patients, because estimation of performance measures in these sites is unreliable [2, 8, 12].

Patient Variables and Risk Adjustment Models

At the patient level, we defined viral control as an HIV RNA ≤400 copies/mL for the last viral load measurement in the year. We used this threshold because in 2009, VHA care systems used varying viral load assays with cutoffs of 50, 75, or 400 copies/mL.

To allow adjustment for case mix, we created a series of variables for patient characteristics that, based on the literature and clinical experience, were known or suspected to be associated with viral control. These included demographics (age, sex, race/ethnicity), CD4 count nadir using VHA laboratory data beginning in 1999, indicators for diagnoses of comorbid conditions (substance use problem, major depression, serious mental illness, or hepatitis C virus infection), patient annual income as a measure of socioeconomic status, and total time on ART in VHA as a measure of treatment experience and potential for viral resistance. Of these variables, only race/ethnicity had missing values (4% missing).

Diagnoses of depression and substance use problems were determined using *ICD-9-CM* codes according to Quan and Elixhauser methods, whereas diagnoses of serious mental illness (schizophrenia, bipolar disorder, or schizoaffective disorder) and hepatitis C were based on validated diagnostic codes and criteria available on the Veterans Aging Cohort Study (VACS) website (www.vacohort.org) [13, 14]. Patients were also considered to have hepatitis C if they had a positive serology in VHA, regardless of follow-up testing to quantify hepatitis C viremia.

We calculated the crude proportion of patients who met the viral control measure in the overall sample, and examined bivariable associations between patient characteristics and viral control at the patient level using χ^2 tests. To examine how case mix varied across care systems, we ranked each care system by unadjusted performance on the viral control measure (see below), and compared patient characteristics by quartile of care system performance.

We then fitted a patient-level multivariable logistic regression model for associations between patient characteristics and viral control, as described in the Supplementary Appendix. This model was used in subsequent analyses that adjusted viral control for case mix using the O/E and RSR methods. The degree of between–care system variation in viral control that remained after adjusting for patient characteristics was summarized by calculating the intraclass correlation coefficient (ICC) [15].

Creation of Report Cards

We made 3 report cards that detailed the performance of each care system on the viral control measure, but which used varying methods. These detailed (1) unadjusted performance, (2) case-mix adjusted performance using the O/E method, and (3) adjusted performance using the CMS RSR method. Report cards included a point estimate of performance for each care system with 95% confidence intervals (CIs). Systems were labeled outliers if their 95% CIs did not overlap the crude overall average for viral control.

The Supplementary Appendix provides the detailed methods used to produce report cards. In brief, the O/E estimator is analogous to the indirectly standardized mortality ratio that is familiar to many HIV clinicians from its use in infectious disease epidemiology [16]. The ratio of the observed number of patients with viral control in each system (O) to the expected number (E) is converted to an estimated performance for the system by multiplying by the population average for viral control. Performance is adjusted for case mix by conditioning the expected number on characteristics of patients in the care system using a fixed-effects logistic regression model. In contrast, the RSR method uses a hierarchical logistic regression model with system-specific random intercepts that represent system quality and account for within-system correlation [12].

To determine the overall influence of individual patient variables on adjusted system rankings, we calculated the Kendall τ correlation coefficient between unadjusted and O/E adjusted system rankings, as described in the Supplementary Appendix. Used in this way, Kendall's τ estimates the proportion of pairs of care systems that change rank order after adjustment, and therefore provides a summary estimate of the overall impact of case-mix adjustment on system rankings [17]. This study was approved by the institutional review board at the University of Iowa and the Iowa City Veterans Affairs Health Care System Research and Development Committee. Analyses were performed using SAS software version 9.2 (Cary, North Carolina).

RESULTS

Overall, 10 913 of 12 368 patients (88.2%) in 91 care systems experienced viral control (Table 1). Care systems were distributed across geographic regions of the United States (Northeast 17, South 38, Midwest 16, and West 20). Compared to patients with viral control, patients without viral control were younger and more likely to be black, to have a nadir CD4 count <100 cells/μL, to have diagnoses of substance use problems, depression, or hepatitis C, and to be on ART for >10 years. They were less likely to have income in the top quartile. With the exception of hepatitis C diagnosis, these associations remained significant in the multivariable model (Table 2), which had moderate discrimination (c statistic = 0.66). After adjusting for patient characteristics, the ICC was 0.09.

Table 1. Patient Characteristics, by Viral Control

Characteristic	Viral Control ≤400 Copies/mL (n = 10 913 [88.2%])	Viral Control >400 Copies/mL (n = 1455 [11.8%])	*P* Value
Age, %			
<30	0.7	1.1	<.01
30–39	5.1	6.1	
40–49	23.2	26.7	
50–65	60.3	60.3	
>65	10.7	5.8	
Sex, male, %,	97.7	97.0	.16
Race/ethnicity, %			
Black	46.9	63.0	<.01
White	39.1	26.8	
Other	9.7	8.0	
Missing	4.3	2.2	
CD4 nadir, cells/μL, %			
0–100	36.0	50.0	<.01
101–200	19.0	21.0	
201–350	23.9	16.8	
351–500	11.1	6.2	
>500	10.0	6.0	
Substance use problem, %	19.6	30.5	<.01
Depression, %	34.0	38.6	<.01
Serious mental illness, %	10.4	11.1	.35
Hepatitis C infection, %	19.4	22.5	<.01
Time on antiretroviral therapy, y, %			
<1	3.7	3.5	<.01
1–5	25.5	23.0	
5–10	30.0	28.0	
>10	40.8	45.5	
Income, $/y, %			
0–3036 (quartile 1)	23.1	25.6	<.01
3037–11 388 (quartile 2)	24.9	28.3	
11 389–25 950 (quartile 3)	25.2	25.8	
>25 950 (quartile 4)	26.8	20.3	

Table 2. Associations Between Patient Characteristics and Viral Control, Multivariable Logistic Regression Model[a]

Characteristic	Odds Ratio (95% CI)
Age, y	
<30	0.33 (.18–.60)
30–39	0.45 (.33–.62)
40–49	0.55 (.43–.71)
50–65	0.67 (.53–.85)
>65	Ref
Race/ethnicity	
Black	0.59 (.52–.68)
White	Ref
Other	0.89 (.72–1.11)
Missing	1.21 (.83–1.78)
CD4 nadir, cells/μL	
0–100	0.51 (.40–.64)
101–200	0.61 (.47–.78)
201–350	0.96 (.74–1.23
351–500	1.16 (.86–1.58)
>500	Ref
Substance use problem	0.65 (.56–.75)
Depression	0.89 (.79–1.00)
Serious mental illness	1.08 (.90–1.30)
Hepatitis C	1.05 (.91–1.21)
Time on ART, y	
<1	1.16 (.84–1.58)
1–5	1.18 (1.01–1.36)
5–10	1.11 (.97–1.23)
>10	Ref
Income, $/y	
0–3036 (quartile 1)	0.77 (.65–.90)
3037–11 388 (quartile 2)	0.77 (.66–.91)
11 389–25 950 (quartile 3)	0.79 (.67–.93)
>25 950 (quartile 4)	Ref

Abbreviations: ART, antiretroviral therapy; CI, confidence interval; Ref, reference.

[a] C statistic = .66; Hosmer-Lemeshow χ^2 7.55 ($df = 8$), $P = .48$.

At the care-system level, performance on the unadjusted viral control measure ranged from 51% to 100% (Figure 1). Compared with patients in systems in the top quartile of unadjusted performance, patients in the bottom quartile systems were more likely to be black and to have a substance use problem (Table 3). System performance was strongly affected by case-mix adjustment, but results differed according to which method was chosen (Figure 1). The RSR method pulled estimated performance for care systems toward the overall mean more strongly than the O/E method (Figure 2). As has been observed in prior studies outside of HIV care, confidence intervals were narrower for the RSR method [18].

Prior to case-mix adjustment, 17 systems were low outliers (ie, poor performers). Following adjustment using the O/E method, 12 systems were labeled low outliers, including 11 from the initial 17. Using the RSR method, 14 were labeled low outliers, including 13 from the initial 17. The O/E and RSR methods agreed on 12 of 14 systems labeled low by either method.

Prior to case-mix adjustment, 11 systems were high outliers. Using the O/E method, 10 were labeled high outliers, including 9 from the initial 11. Using the RSR method, 12 were high outliers, including 7 from the initial 11. The O/E and RSR methods agreed on 8 of 13 systems labeled high by either method. Thus, overall, O/E and RSR agreed on only 20 of 27 (74%) outliers. Systems for which the methods agreed on outlier status had larger patient volumes (median patient volume, 194) compared with systems for which the methods disagreed (median patient volume, 112), but this was not statistically significant given the small sample of outliers ($P = .21$).

Based on Kendall τ correlation values, the patient variables that most strongly affected overall system rankings were, in order of influence, (1) race/ethnicity, (2) CD4 nadir, (3) age, and (4) substance abuse (see Supplementary Appendix for full results). In contrast, removal of depression, serious mental illness, hepatitis C, or time on ART from the adjustment model had less effect on overall system rankings.

DISCUSSION

The performance of HIV care systems on the viral control quality measure was strongly affected by case-mix adjustment, but results differed depending on which statistical method was chosen. Prior to moving forward with public reporting of HIV quality measures, the HIV clinical, research, and policy communities should initiate a public discussion to achieve consensus on best practices for public reporting. This discussion should focus on (1) optimal methods for measuring patient characteristics for use in case-mix adjustment and (2) choice of statistical methods.

Our case-mix adjustment model used only patient variables that were present in administrative data and searchable fields in electronic health records, and had moderate discrimination for viral control (c statistic = 0.66). Although the c statistic does not summarize the overall accuracy of quality report cards, this indicates potential for unmeasured and unadjusted risk related to case mix in our analyses [19]. To accurately emulate the data that are likely to be available to public reporting agencies, we did not adjust for potentially important patient characteristics that are not routinely available in administrative data, such as homelessness, food insecurity, educational status, and attitudes about effectiveness of ART [20–22]. If these patient characteristics both strongly influence viral control and vary in frequency across care systems, then their inclusion in case-mix adjustment models may be necessary. In addition, our comorbidity measures were based on *ICD-9-CM* codes, which are often

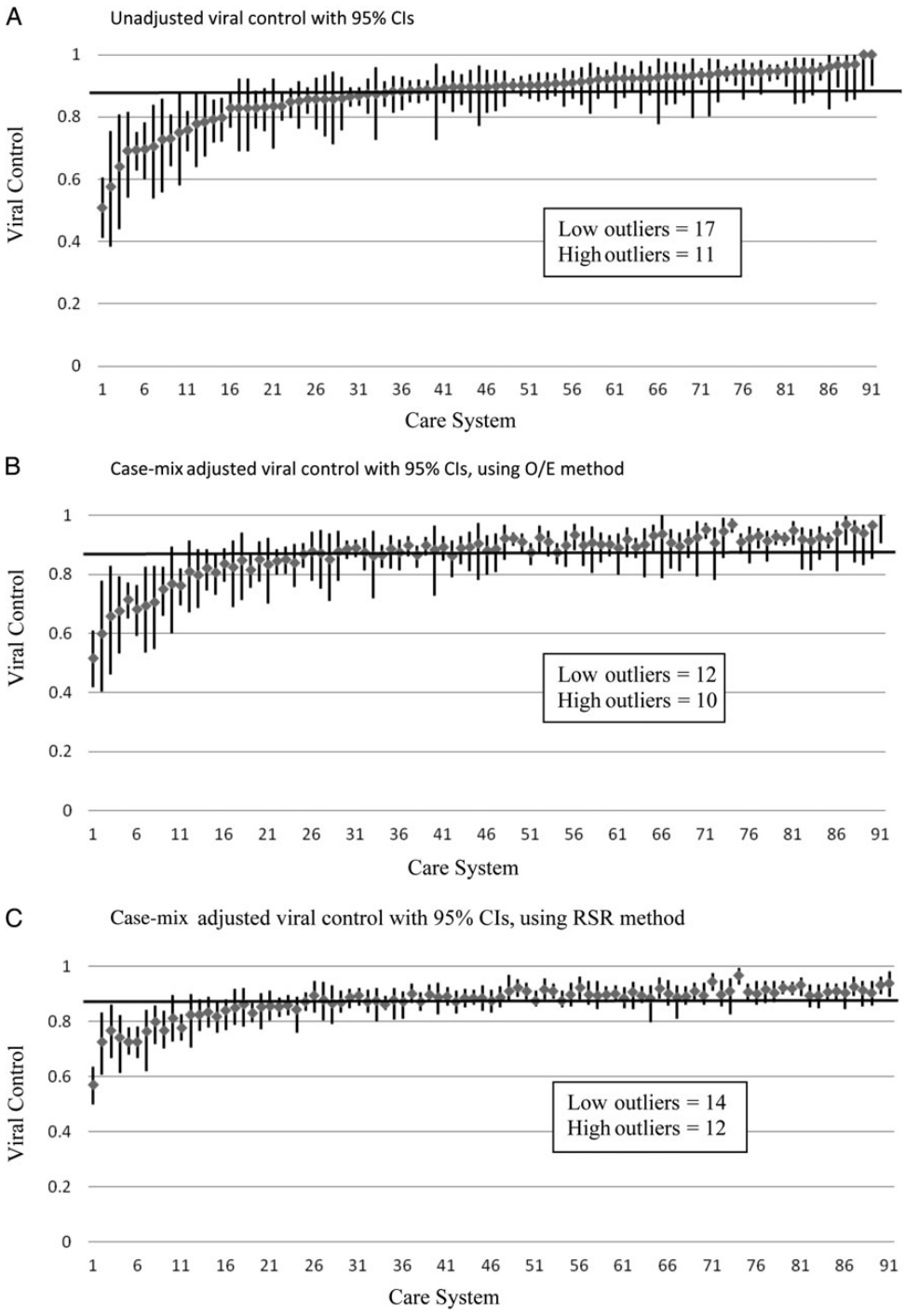

Figure 1. *A–C*, In each viral control report card, the 91 care systems are ordered according to performance on the unadjusted viral control measure. The dark line marks the overall mean for viral control (88%). If the 95% confidence interval for a system does not include the overall mean, the system is labeled an outlier (see "Methods" section). Abbreviations: CI, confidence interval; O/E, observed over expected; RSR, risk-standardized ratio.

imprecise when compared to chart review or other diagnostic instruments [14]. Imprecise measurement of patient comorbidity may lead to residual confounding due to case mix in adjustment models.

Future studies should evaluate whether inclusion of more detailed patient data (eg, measures of homelessness, education, and food insecurity) in adjustment models significantly alters report card results, compared with adjustment using only

Table 3. Patient Characteristics, by Quartile of System Performance on Unadjusted Viral Control Measure

Characteristic	Bottom Quartile (23 Sites, 2307 Patients)	2nd Quartile (23 Sites, 3848 Patients)	3rd Quartile (23 Sites, 3557 Patients)	Top Quartile (22 Sites, 2656 Patients)	*P* Value
Age, y, %					
<30	0.8	0.8	0.8	0.8	<.01
30–39	4.6	4.8	5.9	5.2	
40–49	25.8	19.9	25.6	24.4	
50–65	59.8	64.8	58.0	57.3	
>65	9.0	9.7	9.7	12.2	
Sex, male, %	97.5	97.4	97.3	98.2	.10
Race/ethnicity, %					
Black	58.6	56.7	45.9	32.8	<.01
White	34.6	29.0	39.0	50.9	
Other	4.9	11.6	8.9	11.3	
Missing	1.9	2.7	6.2	5.0	
CD4 nadir, cells/μL, %					
0–100	39.2	36.8	34.8	41.3	<.01
101–200	19.1	19.2	20.8	17.2	
201–350	21.1	24.7	23.2	22.0	
351–500	9.9	10.0	11.2	11.5	
>500	10.7	9.3	10.0	8.0	
Substance use problem, %	22.7	24.4	18.5	17.5	<.01
Depression, %	35.1	34.6	32.5	36.6	.01
Serious mental illness, %	9.7	9.8	11.8	10.3	.02
Hepatitis C, %	19.4	24.8	15.2	19.0	<.01
Time on ART, y, %					
<1	3.9	3.4	3.7	3.8	.01
1–5	25.6	25.5	25.9	23.4	
5–10	30.1	27.6	30.7	31.5	
>10	40.4	43.5	39.7	41.3	
Income, $/y, %					
0–3036	22.3	22.6	26.2	21.7	<.01
3037–11 388	25.8	26.5	24.4	24.3	
11 389–25 950	26.8	24.6	24.5	26.0	
>25 950	25.1	26.3	24.9	28.0	

Abbreviation: ART, antiretroviral therapy.

readily available administrative data. These more detailed patient characteristics could be measured across a large number of care systems within research networks such as the HIV Research Network or VACS [23, 24]. If accurate case-mix adjustment is found to require these more detailed clinical data, then there should be a discussion of the relative costs and benefits of routinely aggregating these data.

The HIV care community should also discuss the appropriateness of including measures of patient race/ethnicity and substance use in risk-adjustment models. On one hand, one may argue that adjustment for these variables "excuses" lower-quality care in systems that serve vulnerable populations with more minorities or higher rates of substance abuse. On the other hand, omission of these variables may unfairly punish specific systems that care for more black persons or persons with substance use problems. This may incent systems to preferentially offer care to individuals who are white and without substance use problems. Such behavior is contrary to the mission of public reporting and could produce significant unintended harm.

Once patient variables are chosen, what is the right method to use when estimating case-mix adjusted performance? This question is difficult to answer due to the absence of a gold-standard measure for true performance for individual systems. Each method has strengths and weaknesses. The O/E method has the advantage of simplicity and familiarity. It is easily interpreted by clinicians and administrators who are familiar with

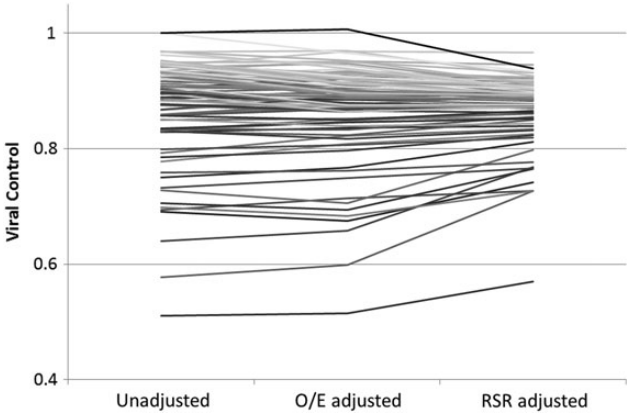

Figure 2. Distributions of viral control for 91 care systems are plotted along the y-axis. Each dark line connects estimated viral control for a single care system, (1) without adjustment for case mix, (2) case-mix adjusted performance using observed over expected (O/E), and (3) case-mix adjusted performance using risk-standardized ratio (RSR). The RSR method pulls estimated performance to the overall mean more strongly than the O/E method (see "Discussion" section). Abbreviations: O/E, observed over expected; RSR, risk-standardized ratio.

fixed-effects logistic regression modeling and indirectly standardized mortality ratios, both of which have long been used in epidemiology [16]. However, the O/E method has been criticized because it does not account for the correlated data structure created by clustering of patients in healthcare systems [6]. It is sensitive to random variation and may falsely label systems as outliers [5]. For these reasons, many reporting programs have moved away from using the O/E method [2].

The RSR method, which is currently used by CMS, addresses these problems by using information from the overall sample to pull performance for care systems to the population average, in proportion to the "noisiness" of system performance. Performance for small systems is most strongly pulled to the mean [6]. As a result, RSR is less likely to falsely label systems—and in particular small systems—as outliers due to random variation [5, 6].

A disadvantage of the RSR method is that, by strongly pulling small sites to the mean, it may obscure underlying associations between care-system size (ie, patient volume) and patient outcomes [9]. If quality is truly lower in low-volume care systems—as has been described in HIV care—then the RSR method may mislead report card users by suppressing volume–outcome associations [9, 25, 26]. Alternative adjustment methods have been proposed that incorporate underlying volume–outcome relationships when profiling the performance of healthcare systems (eg, hierarchical Poisson and Dimick-Staiger estimators), but these are complex and may be difficult to interpret for many report card users [6].

Prior to moving forward with HIV quality reporting, there is need for additional analyses to determine which case-mix adjustment methods are best suited to the specific characteristics of HIV care. Specifically, the potential for inaccurate reporting of HIV quality measures related to underlying volume–outcome relationships and unmeasured patient risk could be evaluated using Monte Carlo simulations and parameter estimates obtained directly from multisite HIV research cohorts.

A limitation of our study is that case-mix adjustment models created in VHA may not apply to or allow comparisons with other care networks, such as Ryan White or Kaiser Permanente clinics. This is true for 3 reasons. First, patient variables aggregated in VHA data warehouses may not be available in other networks. Second, there may be differences in how patient variables are extracted from electronic records and claims data; this could influence how specific variables affect case-mix adjustment. Third, there are important demographic differences between veterans with HIV, who are mostly men, and the overall population of persons with HIV in the United States. In general, it is not valid to directly compare the performance of individual care systems that have nonoverlapping case-mix distributions [27]. For these reasons, it may be appropriate to create separate report cards for healthcare networks with dramatically different data structures or case mix.

CONCLUSIONS

The results of HIV quality report cards are affected by case-mix adjustment, but results differ across statistical methods. Moreover, there is potential for unadjusted risk when profiling HIV care systems using only readily available administrative data to capture case mix. Prior to moving forward with public reporting, the HIV care community should determine optimal methods for measuring case mix and estimating risk-adjusted performance.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online (http://cid.oxfordjournals.org). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Disclaimer. The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs or the US government.

Financial support. This work was supported by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development program (HSR&D). M. E. O. is recipient of a career development award from the HSR&D.

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Horberg MA, Aberg JA, Cheever LW, Renner P, O'Brien KE, Asch SM. Development of national and multiagency HIV care quality measures. Clin Infect Dis **2010**; 51:732–8.
2. Center for Medicare and Medicaid Services (CMS). Hospital Compare. Available at: http://www.medicare.gov/hospitalcompare. Accessed 24 February 2014.
3. Madge S, Smith CJ, Lampe F, et al. An audit of viral load in one clinical population to describe features of viraemic patients on antiretroviral therapy. HIV Med. **2008**; 9:208–13.
4. Pence BW, Miller WC, Gaynes BN, Eron JJ Jr. Psychiatric illness and virologic response in patients initiating highly active antiretroviral therapy. J Acquir Immune Defic Syndr **2007**; 44:159–66.
5. Austin PC, Alter DA, Tu JV. The use of fixed- and random-effects models for classifying hospitals as mortality outliers: a Monte Carlo assessment. Med Decis Making **2003**; 23:526–39.
6. Ryan A, Burgess J, Strawderman R, Dimick J. What is the best way to estimate hospital quality outcomes? A simulation approach. Health Serv Res **2012**; 47:1699–718.
7. Dimick JB, Staiger DO, Birkmeyer JD. Ranking hospitals on surgical mortality: the importance of reliability adjustment. Health Serv Res **2010**; 45(6 pt 1):1614–29.
8. Krumholz HM, Brindis RG, Brush JE, et al. Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. Circulation **2006**; 113:456–62.
9. Silber JH, Rosenbaum PR, Brachet TJ, et al. The Hospital Compare mortality model and the volume-outcome relationship. Health Serv Res **2010**; 45(5 pt 1):1148–67.
10. Backus LI, Gavrilov S, Loomis TP, et al. Clinical Case Registries: simultaneous local and national disease registries for population quality management. J Am Med Inform Assoc **2009**; 16:775–83.
11. Burack JH, Impellizzeri P, Homel P, Cunningham JN Jr. Public reporting of surgical mortality: a survey of New York State cardiothoracic surgeons. Ann Thorac Surg **1999**; 68:1195–200; discussion 201–2.
12. Krumholz HN, Galusha T, Mattera J, Rich A, Wang Y. Risk-adjustment models for AMI and HF 30-day mortality: methodology. Available at: https://www.qualitynet.org. Accessed 24 February 2014.
13. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care **1998**; 36:8–27.
14. Justice AC, Lasky E, McGinnis KA, et al. Medical disease and alcohol use among veterans with human immunodeficiency infection: a comparison of disease measurement strategies. Med Care **2006**; 44(8 suppl 2):S52–60.
15. Snijders T, Bosker R. Multilevel analysis: an introduction to basic and advanced multilevel modeling. London: Sage Publications, **1999**.
16. Rothman K, Greenland S, Lash T. Modern epidemiology. Philadelphia: Lippincott, Williams, and Wilkins, **2008**.
17. O'Malley AJ, Zaslavsky AM, Elliott MN, Zaborski L, Cleary PD. Case-mix adjustment of the CAHPS Hospital Survey. Health Serv Res **2005**; 40(6 pt 2):2162–81.
18. Shahian DM, Torchiana DF, Shemin RJ, Rawn JD, Normand SL. Massachusetts cardiac surgery report card: implications of statistical methodology. Ann Thorac Surg **2005**; 80:2106–13.
19. Austin PC, Reeves MJ. The relationship between the c-statistic of a risk-adjustment model and the accuracy of hospital report cards: a Monte Carlo Study. Med Care **2013**; 51:275–84.
20. Cooper V, Moyle GJ, Fisher M, et al. Beliefs about antiretroviral therapy, treatment adherence and quality of life in a 48-week randomised study of continuation of zidovudine/lamivudine or switch to tenofovir DF/emtricitabine, each with efavirenz. AIDS Care **2011**; 23:705–13.
21. Muthulingam D, Chin J, Hsu L, Scheer S, Schwarcz S. Disparities in engagement in care and viral suppression among persons with HIV. J Acquir Immune Defic Syndr **2013**; 63:112–9.
22. Wang EA, McGinnis KA, Fiellin DA, et al. Food insecurity is associated with poor virologic response among HIV-infected patients receiving antiretroviral medications. J Gen Int Med **2011**; 26:1012–8.
23. Justice AC, Dombrowski E, Conigliaro J, et al. Veterans Aging Cohort Study (VACS): overview and description. Med Care **2006**; 44(8 suppl 2):S13–24.
24. Yehia BR, Gebo KA, Hicks PB, et al. Structures of care in the clinics of the HIV Research Network. AIDS Patient Care and STDs **2008**; 22:1007–13.
25. Kitahata MM, Van Rompaey SE, Shields AW. Physician experience in the care of HIV-infected persons is associated with earlier adoption of new antiretroviral therapy. J Acquir Immune Defic Syndr **2000**; 24:106–14.
26. Wilson IB, Landon BE, Ding L, et al. A national study of the relationship of care site HIV specialization to early adoption of highly active antiretroviral therapy. Med Care **2005**; 43:12–20.
27. Shahian DM, Normand SL. Comparison of "risk-adjusted" hospital outcomes. Circulation **2008**; 117:1955–63.