10 Kaplan GA, Pamuk ER, Lynch JW, Cohen RD, Balfour JL. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *BMJ* 1996;312:999-1003.
11 Pickett KE, Mookherjee J, Wilkinson RG. Adolescent birth rates, total homicides, and income inequality in rich countries. *Am J Public Health* 2005;95:1181-3.
12 US Census Bureau. *Gini ratios by state, 1969, 1979, 1989, 1999*. Washington, DC: US Census Bureau, 1999 (table S4).
13 US Census Bureau. *Census 2000 summary file 3*. Washington, DC: US Census Bureau, 2005.
14 Kochanek KD, Murphy SL, Anderson RN, Scott C. Deaths: final data for 2002. *Natl Vital Stat Rep* 2004;53:1-115.
15 Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Menacker F, Munson ML. Births: final data for 2002. *Natl Vital Stat Rep* 2003;52:1-113.
16 Child and Adolescent Health Measurement Initiative. *National survey of children's health, data resource center on child and adolescent health*. 2005. www.childhealthdata.org.
17 Bradshaw J, Richardson D, Ritakallio V-M. Child poverty and child well-being in Europe. *J Child Serv* 2007;2:18-36.
18 Marmot M, Wilkinson RG. Psychosocial and material pathways in the relation between income and health: a response to Lynch et al. *BMJ* 2001;322:1233-6.
19 Wilkinson RG. Health inequalities: relative or absolute material standards? *BMJ* 1997;314:591-5.
20 Ram R. Further examination of the cross-country association between income inequality and population health. *Soc Sci Med* 2006;62:779-91.
21 Larsen LJ. *The foreign-born population in the United States: 2003. Current population reports*. Washington, DC: US Census Bureau, 2004.
22 *Summary of population statistics 1960-2006*. Stockholm: Statistics Sweden, 2007.
23 Baily MN, Garber AM. *Health care productivity.Brooking papers on economic activity—microeconomics*. Washington, DC: Brookings Institution, 1997:143-215.
24 Dunn JR, Burgess B, Ross NA. Income distribution, public services expenditures, and all cause mortality in US States. *J Epidemiol Community Health* 2005;59:768-74.

Accepted: 26 September 2007

Correlations among measures of quality in HIV care in the United States: cross sectional study

Ira B Wilson,[1] Bruce E Landon,[2] Peter V Marsden,[3] Lisa R Hirschhorn,[4] Keith McInnes,[2] Lin Ding,[2] Paul D Cleary[5]

EDITORIAL by Agins and Holden

[1]Department of Medicine, Institute for Clinical Research and Health Care Policy, Tufts-New England Medical Center, Boston, MA 02111, USA
[2]Health Care Policy and Medicine, Department of Health Care Policy, Harvard Medical School, Boston, MA 02115
[3]Department of Sociology, 630 William James Hall, Cambridge, MA 02138
[4]Harvard Medical School Division of AIDS, Landmark Center 2 East, Boston, MA, 02215
[5]School of Public Health, Yale University, New Haven, CT 06520-8034
Correspondence to: I B Wilson
iwilson@tufts-nemc.org

BMJ 2007;335:1085-8
doi:10.1136/bmj.39364.520278.55

This article is an abridged version of a paper that was published on bmj.com on 29 October 2007. Cite this version as: *BMJ* 29 October 2007, doi:10.1136/bmj.39364.520278.55 (abridged text, in print: *BMJ* 2007;335:1085-8).

ABSTRACT

Objective To determine whether a selected set of indicators can represent a single overall quality construct.
Design Cross sectional study of data abstracted during an evaluation of an initiative to improve quality of care for people with HIV.
Setting 69 sites in 30 states.
Data sources Medical records of 9020 patients.
Main outcome measures Adjusted performance rates at site level for eight measures of quality of care specific to HIV and a site level summary performance score (the number of measures for which the site was in the top quarter of the distribution).
Results Of 28 site level correlations between measures, two were greater than 0.40, two were between 0.30 and 0.39, four were between 0.20 and 0.29, and the 20 remaining were all less than 0.20. One site was in the top quarter for seven measures, but no sites were in the top quarter for six or eight of the measures. Across the eight quality measures, sites were in the top quarter no more often than predicted by a chance (binomial) distribution.
Conclusions The quality suggested by one measured indicator cannot necessarily be generalised to unmeasured indicators, even if this might be expected for clinical or other reasons.

INTRODUCTION

Publicly reported performance data are increasingly available for health plans, hospitals, nursing homes, and physicians, and many providers are now being rewarded on the basis of measures of quality of care.[1-3] These initiatives generally rely on a small set of measures, usually of processes of care but sometimes of outcomes. Common indicators of performance of health plans and physicians focus on the provision of preventive services and the management of a small number of chronic conditions, such as diabetes and asthma.

A rationale for using a small subset of quality indicators is the belief that an organisation's performance on unmeasured processes or outcomes will be similar to that on measured ones. An extension of this logic is that monitoring care indicators for a carefully selected set of prevalent and important conditions, such as diabetes, hypertension, and heart attacks, provides valid information about the overall quality of care provided by a physician, medical group, health plan, or hospital.

Several studies have examined the relations among quality measures for various different types of organisations, but few of these studies examined outpatient medical practices. A recent study of 11 outpatient practices that assessed measures of technical clinical quality found no significant correlations between these measures.[4] Other studies that examined hospitals,[5 6] health plans,[7] and communities[8] have found low correlations among quality measures.

Examining correlations of relations among quality indicators is critical for both measurement and improvement of quality. It is important to understand whether it is appropriate to draw conclusions about the overall quality on the basis of a limited set of indicators. In addition, finding strong correlations among quality measures would support the theory that the measures are the output of a single functional system and that efforts to improve quality should focus on characteristics of the system.

We examined the relation among eight quality indicators for a single chronic condition in care settings in which we expected relatively high correlations—organisations that deliver HIV care to outpatients.

METHODS

Data for this study were collected as part of an evaluation of a quality improvement collaborative that included HIV care.[9 10]

Participants

Site selection—Of the 200 relevant sites in the United States in May 2000, 171 were eligible to participate in the study. We enrolled 44 sites that were participating in the quality improvement intervention and an additional 25 sites that served as controls, giving 69 participating sites in 30 states. We surveyed medical directors at each site to determine specific characteristics.

Patient selection—We randomly sampled 75 active patients from each site before the intervention and a second random sample of 75 after the intervention. The intervention took place from 30 June 2000 to 31 December 2001. Patients were considered active if they had one or more visits during the review period.

Data collection

We collected data from the medical records of each sampled patient over one year of care for the two review periods. Data abstracted included age, sex, history of HIV related illnesses, comorbid medical or psychiatric conditions, screening and prophylaxis for HIV related conditions, number and timing of visits, CD4 cell counts, viral loads, and antiretroviral medications. The first review covered the year before the intervention, and the second covered the year beginning six months after the start of the intervention and ending three months after the end of the intervention.

Quality of care measures

Our primary measures were proportion of use of highly active antiretroviral therapy (HAART) at the time of the last visit during the review period and control of HIV viral load for appropriate patients. Viral load was considered as controlled if it was undetectable or if the total viral load was less than 400 copies/ml. We also assessed whether screening for tuberculosis, hepatitis C, and cervical cancer, appropriate prophylaxis against *Pneumocystis carinii* pneumonia, and influenza vaccinations were provided during the review period. We defined appropriate access to outpatient care as having a visit during at least three of four quarters.

Analyses

Our unit of analysis was the care site. The site level value for each measure of quality was the proportion of patients for whom the quality indicator was documented in the reviewed medical records. After removing duplicated reviews our analyses included a total of 9020 unique patients in care at 69 sites. Because characteristics of patients vary among sites, we examined adjusted means for each of the eight quality measures. We adjusted for patients' characteristics that might be related to the quality measures. Next, we calculated the correlations among the eight adjusted quality measures, and, finally, examined the degree to which high performance on one indicator was related to high performance on the other seven. We dichotomised each quality measure at the 75th centile of the distribution across all 69 sites, and called those sites in the top quarter "high performers." We examined the distribution of the number of times sites performed in the top quarter across the eight quality measures and compared it to a binomial distribution for eight independent trials with a probability of success (being a high performer) of 0.25.

RESULTS

Characteristics of patients and sites—Thirty two per cent of patients were female, 16% reported active substance abuse, and 32% had CD4 cell counts below 200×10^6/l. The 69 sites studied were in 30 states in all regions of the US, representing the full spectrum of types of organisation that provide HIV care. Most sites described themselves as having HIV expertise. Only 3% were general medicine clinics with no specialised HIV team.

Site level quality measures—Clinic performance on the quality measures ranged from 0.38 of patients with non-detectable viral loads (table 1) to 0.81 of eligible patients on HAART and 0.81 of patients with documented hepatitis C status. The greatest variation across clinics was for the proportion of patients who received tuberculosis screening (interquartile range of 0.35-0.69), and the least variation was seen for the proportion of eligible patients who received HAART (interquartile range of 0.77-0.86).

Correlations among quality measures—Of the 28 correlations between measures at the clinic level (table 2), the highest was the relation between proportions of HAART therapy and *P carinii* pneumonia prophylaxis at 0.42 ($P<0.001$). The correlation between the proportion receiving cervical cancer

Table 1 | Measures of quality (proportions, adjusted for patients' characteristics) at the 69 sites studied

Quality measure	Mean (range)	Median (IQR*)
HAART use	0.81 (0.57-0.93)	0.81 (0.77-0.86)
Non-detectable HIV viral load	0.38 (0.01-0.60)	0.41 (0.30-0.48)
P carinii pneumonia prophylaxis	0.70 (0.25-1.00)	0.75 (0.61-0.84)
Tuberculosis screening	0.52 (0.06-0.91)	0.53 (0.35-0.69)
Hepatitis C screening	0.81 (0.25-1.00)	0.87 (0.73-0.93)
Cervical cancer screening	0.60 (0.27-0.98)	0.62 (0.50-0.71)
Influenza vaccination	0.51 (0.02-0.82)	0.53 (0.45-0.61)
Visits in three quarters	0.67 (0.47-0.82)	0.68 (0.62-0.75)

*Interquartile range.

screening and tuberculosis screening was nearly as high at 0.40 (P<0.001). Two other correlations were greater than 0.30, those between the proportion receiving hepatitis C screening and tuberculosis screening (0.32, P<0.01), and between influenza vaccination and non-detectable viral loads (0.30, P<0.05). Four additional correlations were between 0.20 and 0.29, and the 20 remaining correlations were all less than 0.20.

Distribution of number of high performance areas—The number of times sites were in the top quarter (a "high performer") for the eight quality measures ranged from none to seven. The actual and expected distribution under an assumption that "high performance" on different measures occurs at random are not statistically different (P=0.49).

DISCUSSION

We found relatively weak associations between the assessed indicators of quality of HIV care. Of the 28 possible correlations between the eight quality measures, only two (7%) were greater than 0.40, two were between 0.30 and 0.39, and 20 (71%) were less than 0.20. Furthermore, there were no more "high performing" organisations than were predicted by chance. Only one site was in the top quarter for seven measures. We expected that their focus on HIV care would lead some of these sites to develop systems and procedures that would positively affect multiple aspects of care. Moreover, guidelines for HIV care had been widely disseminated when these data were collected.[11-13]

We thought that in HIV care sites the preconditions would exist for high correlations among measures, including focus on a single condition, relatively high proportions of specialisation, and the presence of multidisciplinary HIV care teams. Our results suggest that specialisation and focus on a specific condition may not be sufficient to produce high quality in multiple aspects of care. Consistency may require the coordination of multiple processes and procedures.

P carinii pneumonia prophylaxis and HAART may have been more highly correlated than most other pairs of measures because the processes that produce these outcomes have several shared elements (that is, both are prescriptions given by physicians, and both are guided by CD4 cell counts).

Adequate coordination among processes is probably more difficult to achieve when quality measures assess care given by different providers at multiple care sites (such as different services in a hospital). One potential reason that studies of quality[14 15] and quality improvement efforts[9 16-18] have yielded less impressive results than many expected may be the difficulty of simultaneously improving and coordinating multiple systems.

One interpretation of these data could be that providers recognise that they cannot provide uniformly high quality care and that they therefore prioritise. If providers know that they have to trade off some goals of care against others this is further proof that performance on one measure might imply little about performance on others.

Study limitations

We examined quality measures that could be assessed by reviewing medical records. All but one were measures of process. Our findings might have differed if we had been able to assess mortality, rates of admission to hospital, appropriate management of opportunistic infections, changes in health status, adherence to medication, or patients' reports about care.[19] We think that measurement of other processes or outcomes, however, would yield even lower correlations. Another limitation of reviewing medical record is that processes may have been completed, but not documented, biasing our estimates downwards.

Finally, we studied patients at clinics receiving specific funding, and our findings may not generalise to other HIV care settings. Because this specific funding goes to rural and urban underserved care sites, our findings may not generalise to sites that care for patients with higher incomes, more education, and better health insurance. To the extent that our study design excludes sites with consistently low quality scores, the correlations that we report are lower than they would be in a broader sample.

Our findings have implications for efforts to monitor quality and improvement. Current policy initiatives that seek to pay physicians for their performance on a selected set of indicators or that create tiers of physicians or hospitals may not improve quality across a broad spectrum of care or conditions. Indeed, such programmes could prompt physicians or organisations

Table 2 | Correlations among adjusted quality measures (n=69)*

	1	2	3	4	5	6	7	8
1 HAART use	1.0	—	—	—	—	—	—	—
2 Non-detectable HIV viral load	0.20	1.0	—	—	—	—	—	—
3 *Pneumocystis carinii* pneumonia prophylaxis	0.42***	0.02	1.0	—	—	—	—	—
4 Tuberculosis screening	0.13	0.02	0.18	1.0	—	—	—	—
5 Hepatitis C screening	−0.02	0.10	0.06	0.32**	1.0	—	—	—
6 Cervical cancer screening	−0.09	0.06	0.14	0.40***	0.05	1.0	—	—
7 Influenza vaccination	−0.002	0.30*	0.12	0.18	0.24*	0.07	1.0	—
8 Visits in three quarters	0.03	0.26*	0.04	0.08	0.009	0.21	0.14	1.0

*P<0.05; **P<0.01; ***P<0.001.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Selected indicators are often used as measures of overall quality of care

Few studies have published correlations between indicators of care quality, and none has done so for outpatient specialty care

WHAT THIS STUDY ADDS

There were low correlations among quality indicators for people with HIV disease

It might be hazardous to generalise from performance on a small number of quality indicators to performance on other indicators that are not measured

to channel efforts into affecting the indicators being assessed to the detriment of other aspects of quality.[20]

Implications

Our results suggest that none of the sites we studied had the kinds of administrative, clinical, and human resources systems in place that are necessary to produce consistently high care quality. Focusing on the improvement and coordination of multiple systems within organisations may be a useful direction to pursue.

We thank Carol Cosenza and Patricia Gallagher of the Center for Survey Research who assisted with instrument development and survey administration and our colleagues at the Health Resources and Services Administration and at the Institute for Healthcare Improvement who participated in and facilitated the EQHIV study.

Contributors: See bmj.com.
Funding: Agency for Healthcare Research and Quality (R-01HS10227), and the Lifespan/Tufts/Brown Center for AIDS Research (grant No P30A142853). IBW was supported in part by a mid-career investigator award in patient oriented research from the National Center for Research Resources (K24 RR020300).
Competing interests: None declared.
Ethical approval: The committee on human studies of Harvard Medical School approved the study protocol.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Rosenthal MB, Landon BE, Normand SL, Frank RG, Epstein AM. Pay for performance in commercial HMOs. *N Engl J Med* 2006;355:1895-902.
2 Rowe JW. Pay-for-performance and accountability: related themes in improving health care. *Ann Intern Med* 2006;145:695-9.
3 Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U, et al. Pay-for-performance programs in family practices in the United Kingdom. *N Engl J Med* 2006;355:375-84.
4 Gandhi TK, Francis EC, Puopolo AL, Burstin HR, Haas JS, Brennan TA. Inconsistent report cards: assessing the comparability of various measures of the quality of ambulatory care. *Med Care* 2002;40:155-65.
5 Jha AK, Li Z, Orav EJ, Epstein AM. Care in US hospitals—the Hospital Quality Alliance program. *N Engl J Med* 2005;353:265-74.
6 Landrum MB, Bronskill SE, Normand SL. Analytic methods for constructing cross-sectional profiles of health care providers. *Health Serv Outcomes Res Methodol* 2000;1:23-47.
7 Schneider EC, Zaslavsky AM, Landon BE, Lied TR, Sheingold S, Cleary PD. National quality monitoring of Medicare health plans: the relationship between enrollees' reports and the quality of clinical care. *Med Care* 2001;39:1313-25.
8 Kerr EA, McGlynn EA, Adams J, Keesey J, Asch SM. Profiling the quality of care in twelve communities: results from the CQI study. *Health Aff (Millwood)* 2004;23:247-56.
9 Landon BE, Wilson IB, McInnes K, Landrum MB, Hirschhorn L, Marsden PV, et al. Effects of a quality improvement collaborative on the outcome of care of patients with HIV infection: the EQHIV study. *Ann Intern Med* 2004;140:887-96.
10 Wilson IB, Landon BE, Hirschhorn LR, McInnes k, Ding L, Marsden PV, et al. Quality of HIV care provided by nurse practitioners, physician assistants, and physicians. *Ann Intern Med* 2005;143:729-36.
11 Carpenter CC, Cooper DA, Fischl MA, Gatell JM, Gazzard BG, Hammer SM, et al. Antiretroviral therapy in adults: updated recommendations of the International AIDS Society-USA Panel. *JAMA* 2000;283:381-90.
12 Yeni PG, Hammer SM, Carpenter CC, Cooper DA, Fischl MA, Gatell JM, et al. Antiretroviral treatment for adult HIV infection in 2002: updated recommendations of the International AIDS Society-USA Panel. *JAMA* 2002;288:222-35.
13 Dybul M, Fauci AS, Bartlett JG, Kaplan JE, Pau AK. Guidelines for using antiretroviral agents among HIV-infected adults and adolescents. *Ann Intern Med* 2002;137:381-433.
14 McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeChristofaro A, et al. The quality of health care delivered to adults in the United States. *N Engl J Med* 2003;348:2635-45.
15 Committee of Quality of Health in America, Institute of Medicine. *To err is human: building a safer health system*. Washington, DC: National Academy Press, 2000.
16 Homer CJ, Forbes P, Horvitz L, Peterson LE, Wypij D, Heinrich P. Impact of a quality improvement program on care and outcomes for children with asthma. *Arch Pediatr Adolesc Med* 2005;159:464-9.
17 Schonlau M, Mangione-Smith R, Chan KS, Keesey J, Rosen M, Louis TA, et al. Evaluation of a quality improvement collaborative in asthma care: does it improve processes and outcomes of care? *Ann Fam Med* 2005;3:200-8.
18 Asch SM, Baker DW, Keesey JW, Broder M, Schonlau M, Rosen M, et al. Does the collaborative model improve care for chronic heart failure? *Med Care* 2005;43:667-75.
19 Wilson IB, Ding L, Hays RD, Shapiro MF, Bozzette SA, Cleary PD. HIV patients' experiences with inpatient and outpatient care: results of a national survey. *Med Care* 2002;40:1149-60.
20 Fisher ES. Paying for performance—risks and recommendations. *N Engl J Med* 2006;355:1845-7.

Accepted: 3 September 2007