

Sexually Transmitted Diseases Program Performance Measures: How Are They Performing?

Thomas A. Peterman, MD, MSc, Daniel R. Newman, MA, Dayne E. Collins, BS,
Sonal R. Doshi, MS, MPH, and Stuart M. Berman, MD, ScM

Background: Performance measures were developed in order to improve the performance of sexually transmitted disease (STD) prevention programs.

Methods: A consultant worked with persons from STD programs and Centers for Disease Control and Prevention to identify possible measures. Measures were pilot tested for feasibility and relevance in several programs, then implemented nationwide in 2004. Data were collated and shared with programs and presented at national meetings. Site visits, webinars, and technical assistance focused on program improvement related to the measures. Reported data were analyzed to see if national performance improved on the activities measured.

Results: Some measures were dropped or revised, and quality of reported data improved over time. There was little evidence that overall program performance improved.

Conclusions: Performance measures are one way to monitor performance, and might contribute to program improvement, but additional efforts are needed to improve performance.

The US Public Health Service was given the responsibility to control venereal diseases during the First World War.¹ After the war, control efforts waned until 1936 when Thomas Parran Jr. was appointed Surgeon General of the US Public Health Service. He launched a major syphilis control program.² A national effort to control gonorrhea was added in the 1970s. Since then, the list of sexually transmitted diseases (STDs) has grown to include chlamydia, herpes simplex virus, hepatitis B, human papillomavirus, HIV, and dozens of others.³ These infections present a variety of challenges and opportunities for individuals, health care providers, and STD control programs.

Although the purview of STD control has expanded, STD control programs, like other government (and private) programs, have been targeted by efforts to improve quality. Recent examples of these efforts include the Reinventing Government movement and the Government Performance and Results Act.⁴ Many STD control programs have seen decreasing budgets since the recession of 2008; they cannot afford interventions that do not work well. The need for quality improvement and the demand for evidence of effectiveness will likely intensify as US health reform efforts bring new funding opportunities for prevention interventions that are shown to be effective.

From the Division of STD Prevention, Centers for Disease Control and Prevention, Atlanta, GA

Correspondence: Thomas Peterman, MD, MSc, Centers for Disease Control and Prevention, Mailstop E-02, 1600 Clifton Road NE, Atlanta, GA 30333. E-mail: tap1@cdc.gov.

Received for publication July 2, 2010, and accepted December 6, 2010.

DOI: 10.1097/OLQ.0b013e31820a8d9b

Copyright © 2011 American Sexually Transmitted Diseases Association

All rights reserved.

Performance measures are one approach to quality improvement. Performance measures are designed to describe the effect of an intervention, measure improvement after some modification, or compare the quality of care delivered by different groups.^{5,6} The Division of STD Prevention began investigating the use of performance measures in 1999. Our purpose in designing performance measures for STD programs was not to use performance as a basis for funding; we expected that measurement would motivate improvements in performance at the program level and thereby improve overall performance at the national level. In addition, the measures would help low performers on a particular measure identify high-performing programs that might assist them in identifying solutions to problems with performance. In this article, we describe our 11 years of experience with performance measures, and include an analysis of data from the past 4 years, to determine if there is evidence that our performance measures efforts have led to improved performance.

METHODS

Developing the Measures

The performance measures effort started in 1999 as discussions with a performance management consultant, members of the National Coalition of STD Directors, and individuals from a variety of STD programs. There were 3 meetings in Atlanta during 2000–2001. A logic model was developed to show program components and identify opportunities for measuring performance (Fig. 1). Dozens of possible performance measures were identified. Some measures, such as those used for partner notification, were easily adapted from measures that had been used by programs for decades. However, some important activities, such as collaborating with HIV prevention programs, could not be adequately assessed by a performance measure. Between 2001 and 2003, candidate measures were refined and pilot tested by 7 STD programs. Pilot testing assessed problems encountered in collecting the data and whether the programs thought the data were helpful.

Our intent was to measure components of STD control in a broad public health context rather than to limit our measures to activities directly implemented by STD program personnel. Therefore, we included measures of approaches that STD programs could influence, such as chlamydia screening in juvenile detention centers, which is usually administered and paid for by the department of corrections.

We wanted measures that encouraged good practices, were unlikely to lead to unintended consequences, and could not easily be “gamed” (improving apparent performance rather than actual performance).⁷ This led to long definitions for some measures. For example, “The number of contacts brought to treatment as new syphilis cases within 7, 14, and 30 calendar days from the day of interview of the index case, per case of Primary and Secondary syphilis.” The 3 time periods were

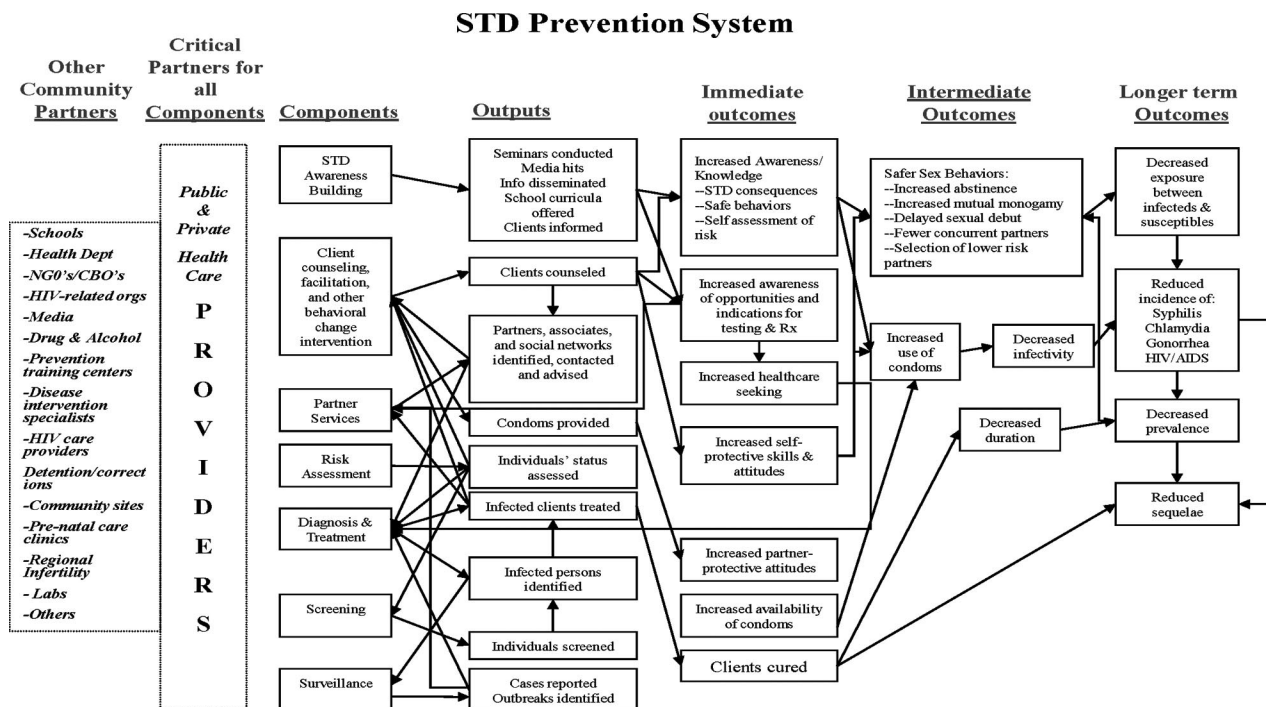


Figure 1. Logic model of an STD Prevention System that was used to develop the performance measures.

meant to encourage rapid notification without discouraging efforts after 2 weeks. We used the number of index cases as the denominator, rather than the number of “partners named,” to encourage interviewing of all index cases and seeking as many partners as possible. The measure could not be gamed by decreasing the denominator, such as listing as “named” only the partners who would be easy to find. Counting only newly diagnosed and treated syphilis cases emphasizes that the main benefit from partner notification comes from treating persons who are infected. A separate measure counts persons who are treated for possible incubating syphilis, another important prevention benefit. It is measured separately because the date of last exposure is difficult to validate, and programs might justifiably differ in who is offered preventive treatment.

Implementation

Performance measurement began in all programs in 2004 with the understanding that the measures would evolve with wider experience. A multidisciplinary performance measures working group monitored the measures. This group included about 10 persons from different professional backgrounds working at Centers for Disease Control and Prevention (CDC) headquarters and 3 senior public health advisors from field sites. This group met periodically with the Division leadership to review findings and discuss challenges. Some data were derived from the surveillance data routinely reported to CDC. Other data were collected twice per year, and entered by the sites into a CDC-developed performance measures database. Sites could use the database to generate reports of their own data and they could generate reports ranking all sites for all measures. At CDC headquarters, data from all sites were reviewed and compared. Senior public health advisors at headquarters were responsible for monitoring the data from their assigned areas. Members of the performance measures working

group worked with the public health advisors to review the data and check data that appeared to be incorrect. Data from all programs were collected and distributed back to all programs in 2006 and 2008. Findings were also presented through numerous webinars and at national meetings.⁸⁻¹⁵ Each year, through 2008, the working group, consultants, and about 10 members from the field met to review the existing measures, the data, and to consider changing measures.

Performance measures site visits were conducted with teams of 3 to 6 persons from headquarters and the field. These visits were initially conducted to learn how STD programs were using performance measures. Later site visits focused on identifying ideas for improving performance. Between September 2006 and April 2008, there were 18 visits that resulted in a total of 247 recommendations for improvement.

Trends in Performance

We assessed the effect of introducing the performance measures on overall program performance by reviewing the performance measures data reported to CDC. Data collection was piloted in 2004. This analysis considers data reported between 2005 and 2008. There were major improvements over time in the quality of the data submitted by the programs, resulting in decreases in the number of corrections requested by data reviewers at CDC. Corrections were made as data errors were identified and there was no systematic accounting of the corrections or types of errors identified; therefore, we cannot report on trends in initial data quality.

We used 2 approaches to compare changes over time for all programs. For the first approach, we restricted the analysis to programs that reported data for at least 5 of the 8 periods, and calculated the mean of all values reported. This approach considers all programs equally important, regardless of size. Next, we confined our analysis to programs that reported data

for all 8 time periods and compared each program's mean reported performance for the first 4 time periods with that program's mean reported performance in the last 4 time periods. The sign test was used to compare the number of programs that improved with the number that worsened over time.

One measure, chlamydia screening in juvenile detention facilities, was studied in more detail to examine the diversity in performance among programs and to see if we could identify factors associated with level of performance or change in performance.

RESULTS

Changes in the Measures

The current measures (Table 1) reflect the evolution that has occurred as problems with data collection, and interpretation were identified. Some of the initial measures were dropped completely. In 2008, we dropped a measure of the "proportions of providers (or clinics) delivering continuing care for >50 HIV-infected individuals, who have protocols for screening those clients for syphilis." This measure might have encouraged some STD programs to establish contacts with key private providers but it was a very crude measure of actual screening, and after the first year, programs did not find the information to be helpful. In 2009, we dropped 2 partner notification measures related to the testing and treating of nonsexual contacts from the social networks of persons with syphilis (suspects and associates). This was an attractive measure to use because programs have collected that data since the 1940s, but in the 2000s, it represented a very small fraction of the work that the programs did.

Some measures were modified to be more comprehensive. Screening measures were changed to collect not only the proportion tested, but also the proportion of those tested who were positive. In 2007, the measure on timeliness of treatment for women diagnosed with gonorrhea or chlamydia in Infertility Prevention Project family planning clinics was expanded to include similar measures for women diagnosed in STD clinics. The measure of timeliness of treatment for primary and secondary syphilis was added in 2009. Also, in 2009, the syphilis partner notification measure began collecting separate reports for persons prophylactically treated, and for persons treated for diagnosed infections. The partner notification measure for gonorrhea (for areas that did not have many syphilis cases) was changed in 2009 to include all gonorrhea cases. Previously, programs were asked about "priority" populations, but priorities varied across different programs and often changed within programs from year to year, making it difficult to compare any 2 data points.

Completeness of Reports

Data reporting improved over time; however, we did not keep records of how often programs were contacted because of questionable data. Problems that were frequently noticed included denominators that were not consistent with data reported through the surveillance system (e.g., number of primary and secondary syphilis cases reported) and denominators that were inconsistently reported from measure to measure. When numerator data were missing (e.g., date treated), programs occasionally also dropped the observation from the denominator, causing an apparent increase in performance. Programs were asked to check on the validity of data that were inconsistent with previous reports.

TABLE 1. STD Performance Measures in 2010, and Measures That Were Started and Dropped

| | |
|----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Screening | Proportion of female admittees to large juvenile detention facilities that were tested for chlamydia (and proportion of females tested that were diagnosed with chlamydia [CT]). |
| | Proportion of female admittees entering selected adult city and county jails that were tested for syphilis (and the proportion of females tested who were newly diagnosed with syphilis, plus the proportion of females diagnosed with syphilis treated within 14 and 30 d of the date of specimen collection). |
| Treatment | Among clients of IPP family planning clinics, the proportion of women with positive CT tests who were treated within 14 and 30 d of the date of specimen collection (similar measures for women tested for gonorrhea [GC] in IPP clinics) (In 2007, we added these measures for women tested in STD clinics). |
| | Proportion of primary and secondary (P and S) syphilis cases treated within 14 and 30 d of the date of specimen collection. |
| Partner notification | Proportion of P and S syphilis cases interviewed within 7, 14, and 30 calendar days from the date of specimen collection, per P and S syphilis case. |
| | No. contacts treated for newly diagnosed syphilis within 7, 14, and 30 calendar days from day of interview of index case, per case of P and S syphilis (same measure for contacts prophylactically treated). |
| | For areas with fewer syphilis cases only: proportion of GC cases interviewed within 7, 14, and 30 d from the date of specimen collection. |
| Surveillance | Proportion of reported cases of GC, CT, P and S syphilis, early latent (EL) syphilis, and congenital syphilis sent to CDC through the National Electronic Telecommunications System for Surveillance (NETSS) that have complete data for age, race, sex, county, and date of specimen collection. |
| | Proportion of reported cases of GC, CT, P and S syphilis, EL syphilis, and congenital syphilis sent to CDC through NETSS within 30 and 60 d from the date of specimen collection. |
| | Proportion of reported cases of P and S syphilis and EL syphilis sent to CDC through NETSS where gender of sex partners is known. |
| Measures that were dropped | Proportions of providers (or clinics) delivering continuing care for <50 HIV-infected individuals, who have protocols for screening those clients for syphilis. |
| | No. suspects and associates tested for syphilis, per case of P and S syphilis. |
| | No. suspects and associates treated for syphilis, per case of P and S syphilis. |

STD indicates sexually transmitted disease; IPP, Infertility Prevention Project; CDC, Centers for Disease Control and Prevention; HIV, human immunodeficiency virus.

Changes in Performance

We found little evidence that performance of STD programs in general improved after the implementation of the performance measures. There was anecdotal evidence that some programs improved, while others got worse. The measures showed no clear overall trend toward improvement.

Screening. Chlamydia screening of young women admitted to juvenile detention facilities increased slightly from a mean of 55% at 70 sites to 58% at 102 sites (Fig. 2). Comparing the first 2 years with the last 2 years, significantly more sites improved (38) than worsened (21) ($P = 0.036$) (Table 2). Data on syphilis screening of women admitted to jails was expected

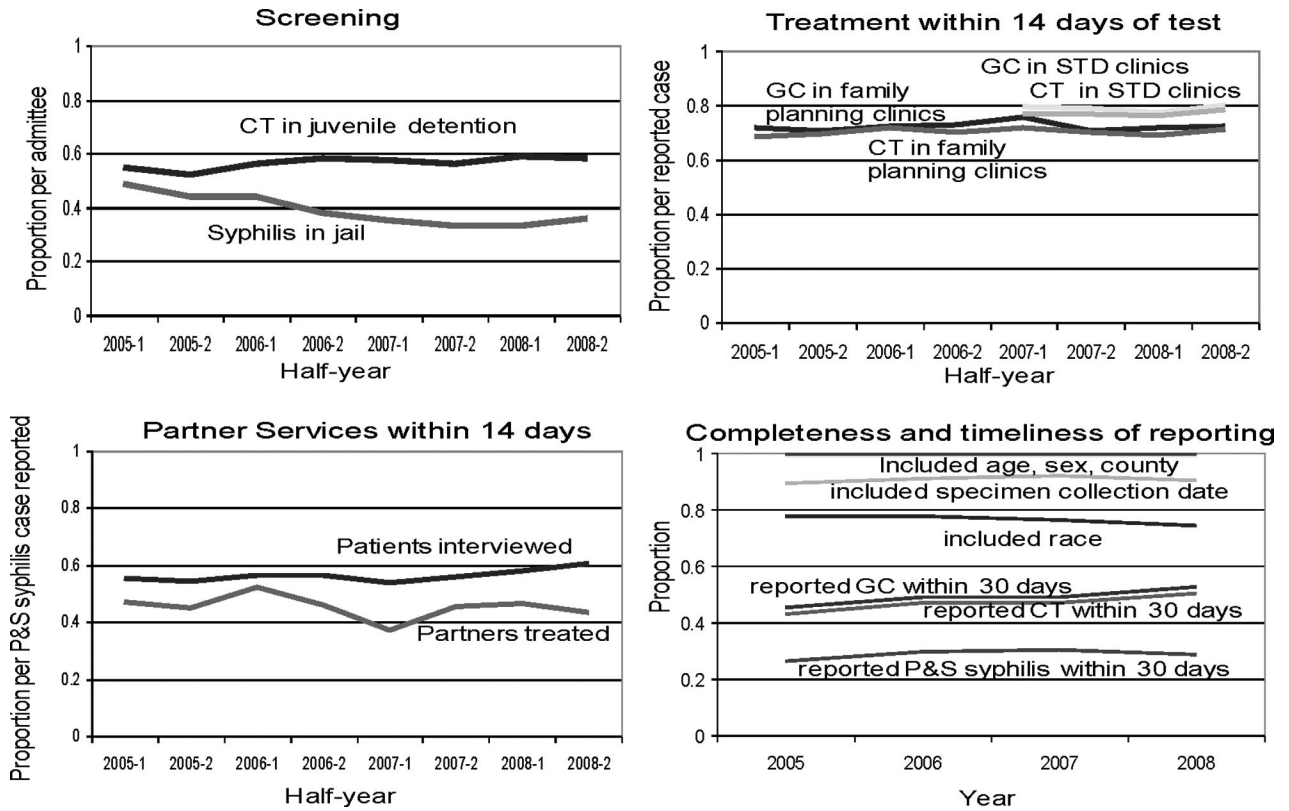


Figure 2. Changes in reported performance of STD programs 2005–2008.

from 30 jails, but reports were received from only 15 jails in the first half of 2005, and 25 jails in the last half of 2008. The mean percent of women tested decreased slightly from 49% at 14 sites to 36% at 19 sites. The number of programs that improved was the same as the number that worsened (6).

Treatment. Treatment within 14 days of specimen collection was measured for gonorrhea and chlamydia in family planning clinics and in STD clinics. In family planning clinics, timely treatment for gonorrhea did not change; 73% at 46 sites, then 74% at 51 sites. Similar numbers of programs improved (18) or worsened (19). Chlamydia treatment in family planning clinics improved slightly from 68% at 46 sites to 72% at 53 sites; similar numbers of programs improved (20) or worsened (21). In STD clinics, timely treatment for gonorrhea was steady at 80% at 47 sites followed by 80% at 49 sites; similar numbers of programs improved (17) or worsened (19). Chlamydia treatment was also similar: 77% at 50 sites and 78% at 49 sites with slightly more improved (23) than worsened (18).

Partner Services. The percent of patients with primary or secondary syphilis who were interviewed about partners within 14 days of the date their specimen was collected increased slightly from 55% at 52 programs to 60% at 56 programs; similar numbers of programs improved (23) or worsened (22). The number of contacts treated within 14 days (prophylactically or for newly diagnosed infection), per primary and secondary syphilis case reported, decreased slightly from 0.47 at 51 programs to 0.43 at 56 programs; slightly fewer programs improved (18) than worsened (27).

Surveillance. The percent of cases reported within 30 days of the date of specimen collection increased slightly for all 3 infections as follows: gonorrhea increased from 45% to 53%, and more programs improved (33) than worsened (22); chlamydia increased from 43% to 50%, and significantly more programs improved (37) than worsened (18) ($P = 0.014$); and syphilis increased from 26% to 29%, and similar numbers of programs improved (25) or worsened (27). Trends in completeness of data were similar for all infections. We show the data for chlamydia. Completeness of data was high (>99%) for age, sex, and county of residence; and for age, significantly more programs improved (34) than worsened (10) ($P = 0.048$). Completeness of data on date of specimen collection was largely unchanged (89%–90%); slightly more programs improved (28) than worsened (21). Completeness of data on race/ethnicity decreased slightly from 78% to 74%; significantly fewer programs improved (21) than worsened (37) ($P = 0.048$).

A more in-depth analysis was done for 1 measure, chlamydia screening in juvenile detention facilities, in order to examine performance of individual programs. The 58 programs were expected to report on all juvenile detention facilities that admitted more than 500 females per year and programs with no large facilities were expected to report on at least one facility from their area. Programs had the option of submitting additional reports from other detention centers. Over the 4 years, there was 1 program that did not report at all. At least 1 report was sent to CDC from 150 facilities. Of these, 104 reported data for at least 4 periods and 60 (from 34 programs) reported data for all 8½-year period.

TABLE 2. Change in Performance, Comparing the Mean From the First 2 Years (2005–2006) With the Mean From the Last 2 Years (2007–2008), for Programs That Reported Data in all Time Periods

| Measure (n) | Improved | Worsened | Unchanged | P |
|--------------------------------------|----------|----------|-----------|--------|
| Screening in juvenile detention (60) | 38 | 21 | 1 | 0.036 |
| Syphilis testing in prison (13) | 6 | 6 | 1 | NS |
| Gonorrhea treatment, IPP (38) | 18 | 19 | 1 | NS |
| Chlamydia treatment, IPP (42) | 20 | 21 | 1 | NS |
| Gonorrhea treatment, STD (39) | 17 | 19 | 3 | NS |
| Chlamydia treatment, STD (42) | 23 | 18 | 1 | NS |
| Syphilis interview (46) | 23 | 22 | 1 | NS |
| Syphilis partner treated (45) | 18 | 27 | 0 | NS |
| Reported within 30 d: | | | | |
| Gonorrhea (58) | 33 | 22 | 3 | NS |
| Chlamydia (58) | 37 | 18 | 3 | 0.014 |
| P and S syphilis (58) | 25 | 27 | 6 | NS |
| CT report included | | | | |
| Age (58) | 34 | 10 | 14 | 0.0004 |
| Sex (58) | 20 | 14 | 24 | NS |
| County (58) | 9 | 8 | 41 | NS |
| Date of specimen (58) | 28 | 21 | 9 | NS |
| Race (58) | 21 | 37 | 0 | 0.048 |

STD indicates sexually transmitted disease; IPP, Infertility Prevention Project; P and S, primary and secondary; NS, not significant.

Two variables were compared—mean performance for the 4 years, and change in mean performance between the first 2 years and the second 2 years (Fig. 3). Five programs averaged more than 98% screened and 6 programs averaged less than

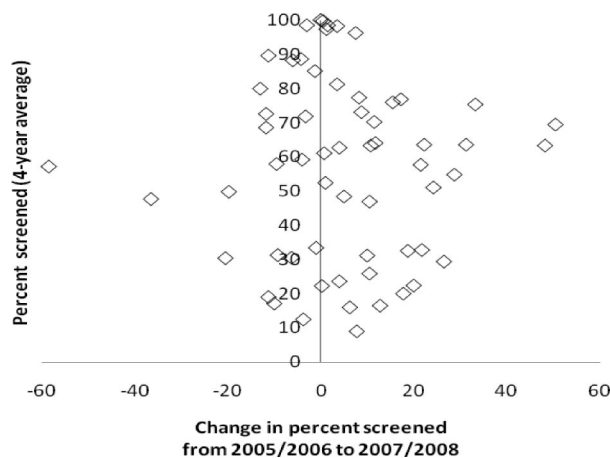


Figure 3. Screening in juvenile detention facilities, average percent screened for 2005–2008, by absolute change in percent screened from the first half (2005–2006) to the second half (2007–2008), for 60 centers that reported data for all 8 time periods.

20% screened; these programs did not have much change in screening over time. Two programs averaged between 60% and 70% screened, but had major increases in screening (from 44% to 95% and from 39% to 88%). One program averaged 58% screened but had a very large decrease in screening (from 87% to 28%). There were no clear differences between these programs with differing performance patterns in terms of size of the facility (average, 52–1796 admittees per year) or geographic location.

DISCUSSION

The performance measures effort was started with the goal of improving STD control. We thought that if we measured and reported programs’ performance on specific activities, then program performance would improve. We expected that the low-performing programs would identify factors that contributed to their low performance and take steps to improve. However, we found little evidence of improvement in performance. Our analysis did not adjust for factors that can influence performance, such as changes in funding or changes in staffing, because we thought those might be changes that programs would make to improve performance. Our intent was to analyze the effect of adding performance measures, and not to study other factors that can influence performance. However, one weakness of our approach is that the performance measures were implemented in all sites at the same time, limiting our ability to assess the potential confounding effects of external factors that could influence performance such as changes in the economy. Programs are currently facing increasing demands with decreasing resources, though our data are from before the state cut-backs due to the 2008 recession.

An in-depth analysis of factors that influence performance on the different measures is beyond the scope of this article, but the diversity of programs’ performance on screening in juvenile detention facilities demonstrates the complexity that underlies overall national performance. Of all the facilities, 24 (41.3%) did not have data for all of the 8 time periods, making it difficult to assess changes in performance. The 60 that did have data for all time periods might be a biased sample if the low-performing sites were less likely to report. Plotting the average performance and change in performance enabled us to identify some distinct groups: consistently high or low performers; and programs that have got much better or worse. The size and location of these facilities were insufficient to explain their screening practices. Further evaluation, including contacting persons responsible for screening, will be needed to determine the factors that enabled individual programs to improve their performance.

Performance measures have been recommended by many,^{16,17} and they have been developed for a variety of health issues. Most are related to health care,^{16,18} though some are more broad, such as maternal and child health or heart disease and stroke prevention.^{19,20} A few reports have assessed the effect of performance measures on the processes that they measure, and the effect in published reports has been generally favorable, but often quite modest.^{17,21–24} Even when there is an impact on process measures, there may not be an effect on outcomes.¹⁷ For example, 4 heart failure performance metrics found substantial improvements (e.g., provision of discharge instructions improved from 21% to 78%), but there was no improvement in 30-day rehospitalization rate or 1-year mortality.²⁵ The effect on health will depend on the effect that the measured activity has on desired outcomes. Some of our performance measures are based on interventions with proven,

quantifiable, outcomes (such as chlamydia screening to prevent pelvic inflammatory disease),^{26,27} but others are not (such as the proportion of gonorrhea cases interviewed).

Good performance measures are valid, quantifiable, influence an important health outcome, and have a reasonable cost. We had many candidate measures that met some, but not all of these criteria. Even the measures that we selected have limited data available to link performance with an outcome. For example, partner notification leads to the identification and treatment of infected persons, but the net effect on STD control has not been measured.²⁸ Our measures have emphasized on syphilis, which was prevalent and devastating when the federal STD program began. However, some programs have very few cases of syphilis, but have many other infections that have no performance measures. There is a tension between spending time developing new measures and working to improve the performance reflected by imperfect existing measures.

Efforts to improve performance of STD control programs face some challenges that are different from those faced by health care providers in other settings. The ability to “pay for performance” is hampered by the variety of activities that are needed in different areas. Because of variations in the populations served, and other conditions beyond the control of local programs, we would not expect that all programs would be equally able to meet the same performance levels. Furthermore, there are no obvious alternatives to relying on health departments for many of the services critical to controlling STD (e.g., surveillance, disease registries, and partner services); therefore, defunding a poorly performing health department does not appear to be an option. We asked programs to set their own performance goals, but some set goals that were easily obtainable, whereas others set goals that were not realistic, and there was no relationship to performance. We are considering the introduction of benchmarks to help programs identify meaningful performance improvement goals. However, setting benchmarks will be difficult, given the variety of populations served by STD control programs.

The changing environment for STD control programs has also been a challenge to programs. In the late 1990s, gonorrhea and syphilis rates were at their lowest levels ever, and rates of chlamydia, a newly reportable infection, were thought to be falling in response to widely implemented screening programs in the Northwest. In the years that followed, many programs became heavily involved in HIV prevention. Remarkable advances in HIV treatment have made it possible to live a near-normal life with HIV, but these advances have been accompanied by treatment optimism, and syphilis epidemics among men who have sex with men.²⁹ Gonorrhea rates have remained relatively stable.²⁹ Chlamydia test positivity rates did not decline, even as screening increased.²⁹ In this environment, perhaps the maintenance of steady performance should be considered a success.

Availability of data for the measures has been a major challenge. Timeliness of treatment was not available for many of the infertility prevention project clinics in the early years. This led to erroneous reports because some sites reported the time of treatment only for cases in which the time of treatment was known. Our measure requires inclusion of all cases reported in the denominator to avoid gaming (by reporting only data from sites with favorable performance). Therefore, our approach will underestimate the proportion treated because it considers cases with missing date of treatment to be untreated.

Quality assurance of the data remains a challenge. Some numbers (such as the number of syphilis cases reported) can be easily checked, whereas others (number of females admitted to jail) cannot be. Our efforts to date have included checking

numbers to be sure that they are internally consistent (the denominator for 7, 14, and 30 days is the same and matches cases reported) and appear to be valid (numbers are fairly consistent from period to period). Programs with data that are outliers are contacted and asked to check their data. We have not done selective audits. There is a trade-off between the value of the data and the cost to obtain it.²⁴

Dissemination of the data has been a challenge. At first, we were concerned that programs might not want their performance data to be seen by other programs, but all agreed that it could be shared. A common database was made available to all programs in 2004, which includes ranking, summary, and trend reports. Programs have been encouraged to use these reports for peer-to-peer program improvement guidance and program planning. In 2006, all data were collated in a detailed report and sent electronically to all programs. In 2007, the performance measures chairperson was detailed to other activities, so the report was never distributed. Data have been reported at conferences and annual meetings during 2007–2010.^{8–15}

Our lack of improvement in measured performance is disappointing, but we believe it is too early to abandon performance measures. Visits to different programs revealed that the performance measures were not accepted by all persons that influence performance, and measures may not even be recognized by some local workers (though some local programs have their own, similar, performance measures). It is unlikely that measurement would influence the performance of persons who are unaware that measurement is taking place. We have relied on the informal incentives of prestige and the satisfaction of good performance which may be effective,⁶ but might also require broader dissemination of results than we have achieved to date. We plan to continue expanding a series of webinars in which programs share best practices. We will continue disseminating performance measures reports and to encourage low performing programs to assess opportunities for improvement. We are providing technical assistance by bringing experts from other programs and from headquarters, to sites that face particularly challenging problems.

The diversity of the tasks involved in STD control, and the diverse areas served by different programs are a challenge for defining and measuring performance. There is little doubt about the programs’ commitment to quality, but it is not clear how to best use performance measures to increase it. We are working to improve the publicity of performance measures data. We are also working to improve buy-in of the performance measures at all levels, training to interpret data, and mechanisms for sharing successful approaches. New approaches to quality improvement (such as plan, do, study, act—PDSA)³⁰ are being introduced to interested programs which might facilitate efforts to improve performance. We will continue to work with programs to help them share approaches that are successful for high-performing programs, and to learn more about challenges facing low performers.

REFERENCES

1. Parascandola J. Sex, Sin, and Science: A History of Syphilis in America. Westport-London, CT-United Kingdom: Praeger, 2008.
2. Parran T Jr. Shadow on the Land. New York, NY: Reynal and Hitchcock, 1937.
3. Holmes KK, Sparling PF, Stamm WE, et al. Introduction and overview. In: Holmes KK, Sparling PF, Stamm WE, et al, eds. Sexually Transmitted Diseases, 4th ed. New York, NY: McGraw Hill, 2008.
4. Kravchuck RS, Schack RW. Designing effective performance-measurement systems under the Government Performance and Results Act of 1993. Public Adm Rev 1996; 56:348–358.

5. Eddy DM. Performance measurement: Problems and solutions. *Health Aff (Millwood)* 1998; 17:7–25.
6. Smith PC. Measuring health system performance. *Eur J Health Econ* 2002; 3:145–148.
7. Wachter RM. Expected and unanticipated consequences of the quality and information technology revolutions. *JAMA* 2006; 295:2780–2783.
8. Parish DD, Peterman TA, Collins D. Assessing timeliness of chlamydia treatment in STD control programs. Presented at: The 17th International Society for STD Research; July 30, 2007; Seattle, WA. Abstract P-081.
9. Brooks LC, Peterman TA, Newman DR, et al. Partner notification interviews for gonorrhea: Who do we target? How many do we find? U.S. 2005–2006. Presented at: The 17th International Society for STD Research; July 30, 2007; Seattle, WA. Abstract P-210.
10. Peterman TA, Newman DA, Collins DE. Cluster investigations of social contacts of persons with syphilis, United States, January 2004–June 2006. Presented at: The 17th International Society for STD Research; July 31, 2007; Seattle, WA. Abstract O-067.
11. Newman DR, Peterman TA, Collins DE. Is time on our side? Timeliness of patient interviews and partner treatment for syphilis, United States, 2004–June 2006. Presented at: The 17th International Society for STD Research; July 31, 2007; Seattle, WA. Abstract P-601.
12. Peterman TA, Newman D, Collins D. Partner notification for syphilis in the United States 2005–2006: Data from the performance measures database. Presented at: The 2008 National STD Prevention Conference; March 11, 2008; Chicago, IL. Abstract P91.
13. Newman DR, Peterman TA, Collins DE. Does monitoring performance influence performance? Data from two performance measures: Timely patient interviews and partner tx for syphilis, U.S., 2005–07. Presented at: The 18th International Society for STD Research; June 29, 2009; London, United Kingdom. Abstract P2.25.
14. Peterman TA, Newman DR, Collins DE. STD program performance measures: Are they performing? Presented at: The 18th International Society for STD Research; June 29, 2009; London, United Kingdom. Abstract P4.102.
15. Newman DR, Peterman TA, Collins DE. Trends in performance: STD program performance measures 2005–2008. Presented at: The 2010 National STD Prevention Conference; March 8–11, 2010; Atlanta, GA. Abstract P61.
16. Leonardi MJ, McGory ML, Ko CY. Quality of care issues in colorectal cancer. *Clin Cancer Res* 2007; 13(suppl 22):6897s–6902s.
17. Humphreys K, Harris AH, Kivlahan DR. Performance monitoring of substance use disorder interventions in the Veterans Health Administration. *Am J Drug Alcohol Abuse* 2009; 35:123–127.
18. Needleman J, Kurtzman ET, Kizer KW. Performance measurement of nursing care: State of the science and the current consensus. *Med Care Res Rev* 2007; 64(suppl 2)10s–43s.
19. Handler A, Grason H, Ruderman M, et al. Assessing capacity and measuring performance in maternal and child health. *Matern Child Health J* 2002; 6:115–126.
20. Croft JB, Dai S, Lawton L, et al. Using outcome measures to monitor the performance of the national heart disease and stroke prevention program: Current capabilities and future challenges. *Am J Prev Med* 2005; 29(5 suppl 1):88–94.
21. Petersen LA, LeChauncy DW, Urech T, et al. Does pay-for-performance improve the quality of health care? *Ann Intern Med* 2006; 145:265–272.
22. Young GJ, Meterko M, Beckman H, et al. Effects of paying physicians based on their relative performance for quality. *J Gen Intern Med* 2007; 22:872–876.
23. Jha AK, Wright SM, Perlin JB. Performance measures, vaccinations, and pneumonia rates among high-risk patients in Veterans Administration health care. *Am J Public Health* 2007; 97:2167–2172.
24. Bratzler DW, Nsa W, Houck PM. Performance measures for pneumonia: Are they valuable, and are process measures adequate? *Curr Opin Infect Dis* 2007; 20:182–189.
25. Fonarow GC, Peterson ED. Heart failure performance measures and outcomes: Real or illusory gains. *JAMA* 2009; 302:792–794.
26. Scholes D, Stergachis A, Heidrich FE, et al. Prevention of pelvic inflammatory disease by screening for cervical chlamydial infection. *N Engl J Med* 1996; 334:1362–1366.
27. Oakshott P, Kerry S, Aghaizu A, et al. Randomised controlled trial of screening for *Chlamydia trachomatis* to prevent pelvic inflammatory disease: The POPI (prevention of pelvic infection) trial. *BMJ* 2010; 340:c1642.
28. Brewer DD. Case-finding effectiveness of partner notification and cluster investigation for sexually transmitted diseases/HIV. *Sex Transm Dis* 2005; 32:78–83.
29. Centers for Disease Control and Prevention. Sexually Transmitted Disease Surveillance, 2008. Atlanta, GA: US Department of Health and Human Services, 2009.
30. van Tiela FH, Elenbaasb TW, Voskuilenc BM, et al. Plan-do-study-act cycles as an instrument for improvement of compliance with infection control measures in care of patients after cardiothoracic surgery. *J Hosp Infect* 2006; 62:64–70.