Periodic Spirometry in Occupational Setting

Improving Quality, Accuracy, and Precision

Eva Hnizdo, PhD, Artak Hakobyan, PhD, James L. Fleming, DO†, and Lu-Ann Beeckman-Wagner, PhD

Objective: Effectiveness of periodic spirometry in medical monitoring depends on spirometry quality. We describe an intervention on spirometry quality and its impact on accuracy and precision of longitudinal measurements. **Methods:** The intervention was conducted from 2005 to 2010 in a monitoring program involving approximately 2500 firefighters. Intervention supported adherence to 2005 American Thoracic Society/European Respiratory Society recommendations through monitoring of spirometry quality and longitudinal data precision, technician training, change of spirometer, and quality control. **Results:** The percentage of forced vital capacity tests meeting the American Thoracic Society/European Respiratory Society criteria increased from 60% to 95% and the mean longitudinal forced expiratory volume in 1 second within-person variation decreased from 6% to 4%. The increased accuracy and precision of measurements and estimated rates of forced expiratory volume in 1 second decline were statistically significant. **Conclusion:** Monitoring of quality and data precision helped to recognize the need for intervention. The intervention improved accuracy and precision of spirometry measurements and their usefulness.

Epidemiologic studies demonstrate that decreased forced expiratory volume in 1 second (FEV_1) and excessive decline in FEV_1 are risk factors for premature death from chronic obstructive pulmonary disease, heart disease, and all-cause mortality.[1] In at-risk occupational settings, periodic spirometry is often conducted to monitor the level of lung function and the change in lung function over time to prevent work-related respiratory injuries, to establish and maintain workers' fitness to wear respirators, or to maintain workers' general health.[2–6] In relatively healthy worker populations, a larger proportion of workers may benefit from the prevention of excessive lung function decline (ie, longitudinal evaluation) than from decision making based on a single abnormal spirometry test.[2,3]

Nevertheless, effective utilization of periodic spirometry at workplace often hinges on having accurate measurements and sufficiently precise longitudinal spirometry data.[2,7,8] Although the expected rate of decline in FEV_1 in healthy adult never-smokers is around 20 to 30 mL/yr,[9] because of the inherent variability in the longitudinal FEV_1 data, it can take five or more years to establish an individual's rate of FEV_1 decline reliably. Nevertheless, it is possible to determine whether a decline over a shorter period (1 to 7 years) is excessive using a limit of longitudinal decline that accounts for FEV_1 data variability, but the reliability of this method still depends on the existing data precision (ie, within-person variation).[2,7]

Recommendations have been provided on how to evaluate and standardize the quality of a single spirometry test[2,10,11] and how to evaluate spirometry quality over time in research settings.[12] Nevertheless, epidemiologic evidence on achieving and maintaining accurate spirometry measurements and acceptable longitudinal data precision in spirometry monitoring is limited. This information is crucial, however, to answer the question on whether the rate of lung function decline can be measured sufficiently accurately in occupational settings to detect early stages of respiratory disease or identify an effect of occupational exposure in a group of individuals and what level of data precision is required.[13]

In this case report, we document the impact of several intervention measures done to improve and maintain acceptable spirometry quality and longitudinal spirometry data precision in an ongoing spirometry-based medical monitoring program conducted in firefighters, that is, an occupational setting with increased risk of work-related respiratory injury.[14] We evaluated the impact of the intervention on the accuracy and precision of the lung function measurements and the estimated rate of lung function decline.

MATERIAL AND METHODS

The intervention on spirometry quality was conducted in an ongoing spirometry-based medical monitoring program that serves approximately 2500 firefighters. Annual spirometry tests have been done since 1988 by several technicians, and the computerized results are utilized by several physicians for disease prevention. Since 2005, Spirometry Longitudinal Data Analysis (SPIROLA) software[15,16] has been used to assist the health care providers in longitudinal spirometry data interpretation. The analysis revealed that the longitudinal spirometry data variability (as measured by the relative pair-wise within-person variation)[7] increased significantly after the year 2000 (Fig. 1), mainly after a new flow-based spirometer was introduced.

The intervention initiated in April, 2005, involved the following measures: technician training according to a National Institute for Occupational Safety and Health (NIOSH)–approved spirometry training course in April, 2005,[17] replacement of the flow-based spirometers (used since July, 2001) with two dry-rolling seal volume spirometers in January, 2006, and monitoring of spirometry data quality and precision since July, 2006. Beginning from 2007, a central quality control program was implemented whereby one of the senior technicians provided monthly feedback to individual technicians regarding their test results' adherence to the 2005 American Thoracic Society (ATS)/European Respiratory Society (ERS) standards and recommendations for improving test quality.[10] In April, 2008, in response to the monitoring results, additional training was provided to improve test quality.

The Occupational Marketing Inc (Houston, TX) (OMI) spirometry software,® integrated with a SensorMedics dry-rolling seal volume spirometer, provided real-time spirometry tracings and generated an automated acceptability error code for each tracing. On the basis of these errors, the system assigns quality grades (A to

From the Division of Respiratory Disease Studies (Drs Hnizdo and Beeckman-Wagner), National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Morgantown, WV; System Research & Applications International, Inc (Dr Hakobyan), Morgantown, WVa; and Phoenix Fire Department (Dr Fleming), Phoenix, Ariz.

This work was supported by funding through the National Institute for Occupational Safety and Health intramural program. Periodic spirometry in occupational setting: improving quality, accuracy, and precision.

†Deceased.

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Mention of a specific product or company does not constitute endorsement by the Centers for Disease Control and Prevention.

Address correspondence to: Eva Hnizdo, PhD, Division of Respiratory Disease Studies, National Institute for Occupational Safety and Health, 1095 Willowdale Rd, Morgantown, WV 26505 (ehnizdo@cdc.gov).

DOI: 10.1097/JOM.0b013e31823078b8

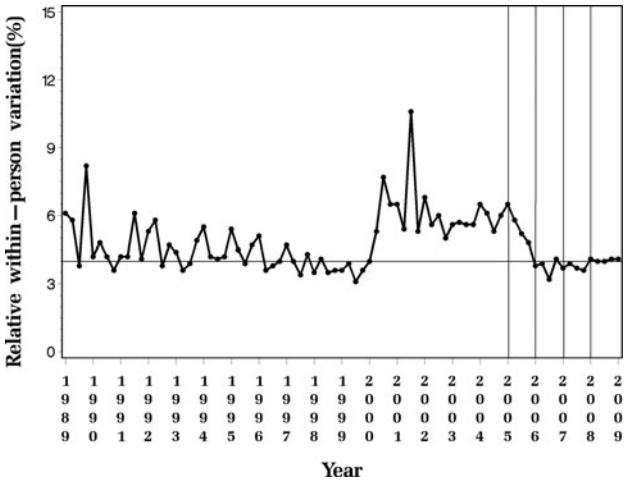

FIGURE 1. Relative pair-wise within-person variation (based on tests repeated annually) for the firefighters' monitoring program. Vertical lines indicate interventions (technician training in April, 2005, volume-based spirometer since January, 2006, onset of central quality control by a senior technician with monthly feedback in 2007, additional training in April, 2008). The goal is to maintain relative within-person variation at 4% or lower.

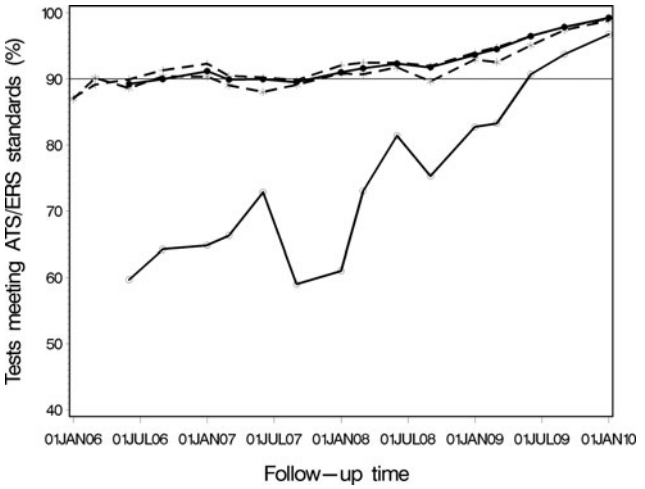

FIGURE 2. Percentage of spirometry tests that met the American Thoracic Society/European Respiratory Society standards for acceptability and repeatability for forced vital capacity (solid line, symbol "o") and for forced expiratory volume in 1 second (solid line, symbol "•") and percentage of tests that met the repeatability criteria for forced vital capacity (dashed line, symbol "*") and for forced expiratory volume in 1 second (dashed line, symbol "+").

D, F) for each testing session to reflect whether the forced vital capacity (FVC) and FEV_1 tests conform to the 2005 ATS/ERS criteria of acceptability and repeatability.[10,11] Nevertheless, the computer-generated quality grades serve as a guide and cannot replace human judgment. If in disagreement, the senior reviewer can overwrite the grade. For example, if the subject has reached a plateau before 6 seconds of exhalation, indicating an acceptable FVC was obtained, the trial is acceptable. SPIROLA software interfaced to the OMI system monitors on a quarterly basis, the percentage of tests that meet the minimum ATS/ERS acceptability and repeatability criteria (two acceptable curves and repeatable values [$\leq$150 mL between the best test and the second best test]; ie, grades A and B). Longitudinal FEV_1 data precision is also monitored using the pair-wise within-person variation, by individual technicians, and overall.

Current smoking status (smoker yes/no), ascertained at the time of the annual medical examination, was collected since 2006 in the OMI and SPIROLA databases.

DATA ANALYSIS

For the analysis, we used deidentified data from the OMI and SPIROLA databases. We calculated the mean and mean 95% confidence limits and *t* test statistics to compare the pre-intervention and post-intervention lung function values. The simple linear regression model was applied to estimate the rate of FEV_1 change (mL/yr) over 1 to 5 years of follow-up in individuals. To evaluate longitudinal spirometry data precision, the pair-wise within-person variation, based on the annual FEV_1 within-person differences, was calculated on an annual basis from the onset of spirometry monitoring in 1988 to 2010.[7,8] The relative within-person variation was standardized using the baseline lung function(s).[7]

The project was approved by the NIOSH Human Subject Review board.

RESULTS

The impact of the interventions was evaluated in terms of the following outcomes: improvement in longitudinal FEV_1 data precision; the percentage of spirometry tests fulfilling the ATS/ERS recommendations; the decline in most frequent spirometry testing errors; and the accuracy and precision of FVC and FEV_1 measurements and of the estimated mean rate of lung function decline.

Longitudinal Data Precision

Figure 1 shows the changes in longitudinal FEV_1 data precision measured as the relative pair-wise within-person variation over the years 1988 through 2009. The precision has been monitored retrospectively and prospectively since 2005 (the value for 2009 was based on within-person differences that included data from 2010). The training in April, 2005 (first horizontal line), resulted in subsequent decrease in the FEV_1 within-person variation, that is, an increase in data precision. Introduction of a volume-based spirometer in January, 2006 (second horizontal line), resulted in further decrease in FEV_1 variation, reaching the value of 4%. Addition of central quality control in 2007 did not substantially decrease the FEV_1 variation, which has been maintained at the 4% level.

Percentage of Tests Fulfilling Current ATS/ERS Recommendations

Figure 2 shows changes in spirometry quality grades, monitored on a quarterly basis as (i) the percentage of tests that fulfilled the ATS/ERS criteria of acceptability and repeatability for FEV_1 and FVC and (ii) the percentage of tests that were repeatable for FEV_1 and FVC (ie, the difference between the two best tests $\leq$150 mL). Initially, around 60% of the FVC and 90% of the FEV_1 measures met the minimal ATS/ERS criteria. Following the additional training in 2008, the percentage gradually increased to above 95% for all measures. The retraining of the technicians in 2008 was focused primarily on improving skills to coach subjects to improve the test quality and to improve performance of the internal quality control reviewer.

Percentage of Trials With Spirometry Testing Errors

The increase in the percentage of FVC tests meeting the ATS/ERS criteria was primarily due to the decline in three types of errors. Figure 3 shows the percentage of trials with various types

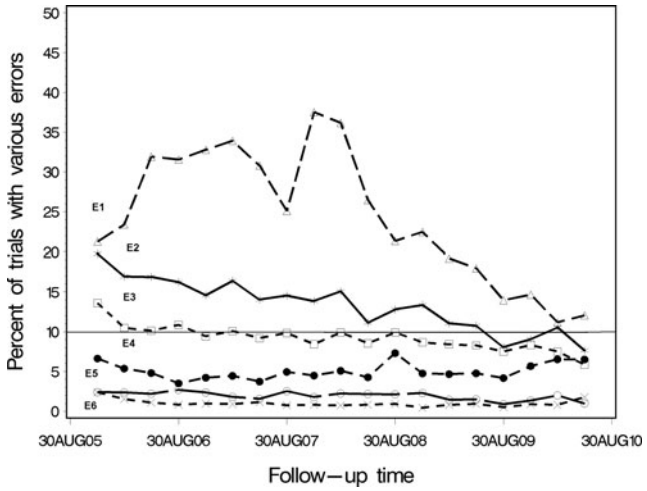

FIGURE 3. Percent of trials with various types of errors. E1 = no plateau was reached (dashed line, triangle), E2 = time to peak expiratory flow is too large (solid line, star), E3 = peak expiratory flow is too low (dashed line, square), E4 = time to exhalation less than 6 seconds (dashed line, dot), E5 = extrapolated volume too large (long dashed line, circle), and E6 = cough during the first second (dashed line, cross).

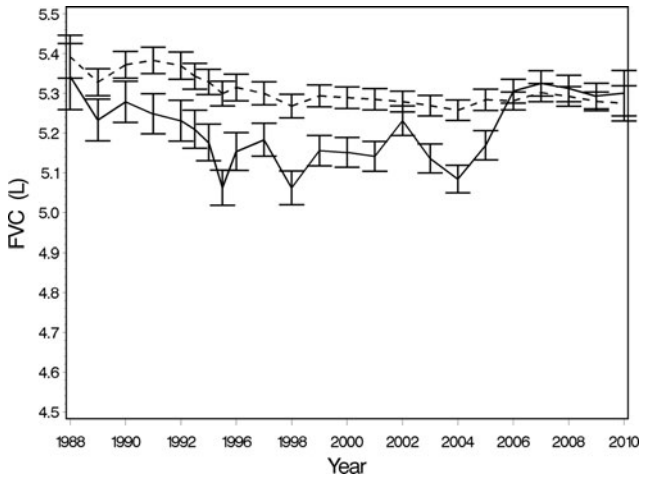

FIGURE 4. Annual mean values and 95% mean confidence interval for observed forced vital capacity (solid line) and predicted forced vital capacity based on the group demographics (age, gender, height, and ethnicity) (dashed line). The intervention on spirometry quality since the year 2005 resulted in significantly higher mean FVC values starting from the year 2006.

of errors as recorded by the OMI system. The most frequent error (E1) was the best volume–time curves not reaching a plateau (initially about 35% declined to 12%). The second most frequent error (E2) was slow time to reach peak expiratory flow (PEF) (initially about 20% declined to 7%). The third most frequent error (E3) was low PEF due to poor effort (initially 14% declined to 6%). Errors that did not change frequency were (E4) less than 6 seconds to exhalation, (E5) cough during first second, and (E6) large extrapolated volume.

Accuracy and Precision of Lung Function Measurements

The improvement in spirometry quality resulted in higher and more accurate measurements of FVC when compared with the predicted values (Fig. 4) and FEV_1. The mean of the measured FVC values (solid line) by year of follow-up became consistently higher than the predicted mean values (estimated using US reference equations[18]) from the year 2006, suggesting higher accuracy and validity of the measurements. The 95% confidence intervals (CIs) became also narrower than for the previous years, indicating higher precision of the observed means. A similar pattern was observed for FEV_1.

To estimate the effect more accurately, Table 1 shows descriptive statistics for individuals who had spirometry measurements during both the most recent pre-intervention years 2003 to 2004 and post-intervention years 2008 to 2009, and compares the mean pre- and post-intervention lung function measurements. (Each person contributed a single test result per period from the most recent test.) The study sample consisted of 84% whites, 4% African Americans, and 12% Hispanics, 5% of the sample were women, and 94.5 reported being nonsmokers during the second period. After an average 4.5 years of follow-up, the mean weight and body mass index (kg/m^2) have increased significantly. For FEV_1, the mean values have not changed despite the 4 years of follow-up and increased body mass index, but the percent predicted values have increased significantly indicating that the accuracy of the measurements has increased. For FVC, both the mean measured values and percent predicted values have increased significantly, demonstrating higher post-intervention accuracy of the FVC measurements. As expected, the ratio decreased with increased accuracy of the FVC measurements.

Precision of the Estimated Mean Rate of Lung Function Decline

It is expected that higher accuracy and precision of individual lung function measurements would lead to more accurate and precise estimates of the rate of lung function decline in individuals and in a group. Figure 5 shows the means (95% mean CI) for the individual linear regression slopes estimated over increasing duration of follow-up, by three periods of data precision observed in Fig. 2, that is, years 1994 to 2000, 2001 to 2005, and 2006 to 2010. For two "precise data" periods when the mean relative within-person variation ranged around 4%, the periods 1994 to 2000 (solid line) and 2006 to 2010 (dashed line), the estimated mean slopes reached the expected rate of -30 mL/yr from 2 years of follow-up and the 95% CIs became relatively narrow, indicating acceptable accuracy and precision of the estimated slopes. On the contrary, for the "imprecise data" period from July, 2001, to January, 2005 (dotted line and 95% CI box), where the mean relative within-person variation ranged around 6%, the estimated mean slopes did not attain the expected value of -30 mL/yr and the 95% CIs were wide. The number of individuals included in the estimation and their mean age at baseline were similar for the three periods: 1362 and 38.8 (SD = 9.4) for the period 1994 to 2000; 2089 and 39.6 (SD = 9.6) for the period 2001 to 2005; and 2735 and 38.8 (SD = 9.9) for the period 2006 to 2010.

DISCUSSION

Achieving and maintaining adequate periodic spirometry data accuracy and precision is important in at-risk occupational settings to identify and prevent work-related respiratory injury, especially to identify individuals who experience excessive decline in lung function. Our study investigated the impact of a multifaceted intervention on improving spirometry quality in a firefighters' spirometry-based medical program. The impact was measured by the following outcomes: longitudinal data precision, percentage of tests fulfilling the ATS/ERS criteria of acceptability and repeatability, and the accuracy

TABLE 1. Descriptive Statistics for 2043 Firefighters Who Had Lung Function Measurements Both Prior to the Onset of the Intervention on Spirometry Quality in 2003–2004 and After Intervention in 2008–2009

	Pre-Intervention Period 2003–2004		Post-Intervention Period 2008–2009		
Characteristics	**Mean**	**Mean 95% CL**	**Mean**	**Mean 95% CL**	***P***
Age (yr)	39.2	38.8–39.6	43.7	43.3–44.1	<0.0001
Height (cm)	179.0	178.7–179.3	179.0	178.7–179.3	0.75
Weight (kg)	92.0	91.4–92.6	93.5	92.9–94.2	<0.001
BMI (kg/m^2)	28.7	28.5–28.8	29.1	29.0–29.2	<0.001
FEV_1 (L)	4.10	4.08–4.14	4.10	4.07–4.13	0.77
FEV_1 (%, predicted)	97.0	96.4–97.5	99.6	99.0–100.1	<0.0001
FVC (L)	5.15	5.11–5.18	5.23	5.19–5.26	<0.005
FVC (%, predicted)	97.2	96.7–97.7	100.4	99.9–100.9	<0.0001
FEV_1/FVC (%)	79.9	79.6–80.2	78.6	78.4–78.9	<0.0001
FEV_1/FVC (%, predicted)	99.5	99.1–99.8	99.0	98.7–99.4	0.06

BMI, body mass index; CL, confidence limit; FEV_1, forced expiratory volume in 1 second; FVC, forced vital capacity.

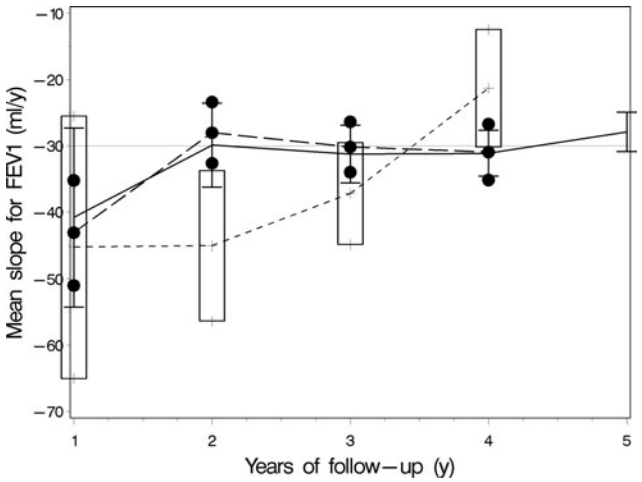

FIGURE 5. Means (95% mean confidence limits) of individual linear regression slopes estimated for increasing follow-up time (1 to 5 years) for two "precise data" periods, that is, within-person variation of 4%, years 1994 to 2000 (solid line, dot symbol) and years 2006 to 2010 (dashed line, dash symbol) and for an "imprecise data" period, that is, within-person variation of 6%, years 2001 to 2005 (dotted line, 95% mean confidence interval as a box).

and precision of the spirometry measurements and of the estimated mean rate of lung function decline. The results provide guidance on what level of longitudinal data precision is adequate and what measures could be considered to improve spirometry quality to achieve the adequate data precision.

To determine whether an individual decline in FEV_1 is accelerated requires sufficiently precise longitudinal data. Our results show that monitoring of longitudinal spirometry data precision (eg, using the pair-wise within-person variation statistic) helps to determine whether the periodic spirometry data are sufficiently precise and whether there are any time-related changes in the data precision that may be related to procedural changes or an effect of occupational exposure, for example, a change of spirometer and associated procedural errors (Fig. 1). Monitoring of the relative, rather than the absolute within-person variation (eg, relative to the baseline lung function), standardizes for the lung function size (mainly because of body size) and in healthy individuals reflects mainly effects of procedural errors.

Furthermore, monitoring of spirometry quality grades provided by some spirometer types helps to identify how well the spirometry quality adheres to the ATS/ERS recommendations for acceptability and repeatability (Fig. 2). Generally, the computer-assigned quality grades are helpful but should be used only as a guide. The senior reviewer should make the final determination of the quality grades and make decisions such as whether to include values, which may have been ignored by the software (eg, a valid FEV_1 may have been deleted because the FVC was not acceptable and vice versa). Monitoring of these quality grades stratified by a technician code further helped to identify procedural errors associated with individual technicians or with the testing equipment.

Poor spirometry quality can be due to technical factors or poor communication between a technician and the subject.[2,19] In our study, technician training was provided in response to the monitoring results at the onset of the intervention in 2005 and again in 2008. The individualized training in 2008 focused on improving technicians' coaching techniques for better subject performance and on improving the performance of the internal quality control reviewer. The training was effective in increasing the percentages of FVC tests reaching the ATS/ERS criteria from 60% to above 95% and also improving quality of FEV_1 (Fig. 3). The most frequent errors reduced during the intervention were as follows: the volume–time curves did not reach a plateau, the time to reach PEF was too large, and low PEF due to poor effort. Generally, tests with these errors are still acceptable, but caution needs to be exercised especially in the longitudinal interpretation. Because spirometry is effort dependent, technicians need to be appropriately trained and motivated to perform the very best on every employee and be also capable of judging the subject's degree of effort and cooperation. Regular refresher training is also recommended to maintain technicians' skills and motivation. The NIOSH-approved spirometry training course has been regarded as a benchmark and best practice, and in 2005, ATS/ERS endorsed the course as a prototype for technician training.[20] National Institute for Occupational Safety and Health–approved courses are listed on the NIOSH Web site.

Inadequate accuracy of the testing equipment or associated procedural errors can be also a source of excessive data variability

as has been demonstrated in this report (Fig. 1) and need to be addressed. For example, the flow-measuring spirometers have the potential for zero flow errors that can affect the end of the test or the measured values.[2] When purchasing a spirometer, health care providers should investigate spirometer's accuracy and capacity for data management and keep in mind that evaluation of lung function changes over time, essential in an effective spirometry monitoring, requires accurate measurements.

Respiratory health of a group can also substantially affect the longitudinal spirometry data variability (ie, group average within-person variation). In a study of early chronic obstructive pulmonary disease, the lung health study, where spirometry quality was strictly monitored,[11] the group mean relative within-person variation in FEV_1 established over a 5-year follow-up period increased with quartiles of baseline bronchial hyperreactivity score from 4.3%, 4.9%, 5.8%, to 7.3%.

The results from our study provide guidance for spirometry monitoring programs onto what levels of data quality and precision are required to obtain accurate and precise measurements of the level of lung function and of the rate of lung function decline. On the basis of our results, in relatively healthy working populations, the level of data precision, as measured by the relative within-person variation, should be at 4% level or lower. It is possible to achieve this level of longitudinal FEV_1 data precision by adhering to the most current ATS/ERS criteria of acceptability and repeatability.[8,21–23] Nevertheless, maintaining this level of data precision requires ongoing monitoring, evaluation, and maintenance of the spirometry quality and data precision to address spirometry quality issues as soon as these arise.

By monitoring the group average relative within-person variation, it is also possible to determine the limit of longitudinal decline appropriate for individuals participating in a particular monitoring program.[7] It needs to be pointed out, however, that monitoring of year-to-year changes has low predictive value of longer-term rate of decline, and that all the longitudinal spirometry data should be considered in the longitudinal evaluation.[7] Annual evaluation can be useful, however, in clinical or occupational settings where the effect of treatment or exposure may be acute and observable within a shorter period.[24] Even if longitudinal assessment is not conducted, lung function measurements need to be accurate. Falsely reduced FVC values result in higher FEV_1/FVC ratio and a failure to recognize airflow obstruction or false-positive restriction. The high mean FVC values in our study are compatible with firefighters' need to maintain their general, as well as respiratory, fitness.

Limitations of the study include lack of occupational exposure, smoking, and symptoms data for the firefighter cohort and inability to evaluate their effect on data variability over time. Nevertheless, the 95% nonsmoking rate in this firefighter's population observed in 2010 is likely in response to the respiratory fitness requirements introduced since the late 1980s when the spirometry monitoring started. Also, there has not been any change in the employment pattern to explain the observed increase and decrease in longitudinal data variability (Fig. 1).

CONCLUSION

In periodic spirometry, it is essential to ensure that the spirometry tests are done according to the most recent ATS/ERS recommendations to obtain accurate and precise lung function measurements and estimates of the rate of lung function decline in individuals and groups. Computerized monitoring of spirometry quality and longitudinal spirometry data precision, and regular oversight by an experienced technician and physician in charge of the program, help in maintaining both spirometry quality and acceptable longitudinal data precision and in the recognition of the need for additional technician training or improvement in the accuracy of a spirometer.

ACKNOWLEDGMENTS

The authors thank Diana Freeland (NIOSH) for providing technician training and quality review oversight and also the medical staff of the Phoenix Fire Department for their cooperation and dedication to implementing all the interventions in this study.

REFERENCES

1. Baughman P, Marott JL, Lange P, Hnizdo E. Health outcomes associated with lung function decline and respiratory symptoms and disease in a community cohort. *COPD: J Chron Obstruct Pulm Dis.* 2011;48:103–113.
2. Townsend MC and the Occupational and Environmental Lung Disorders Committee. Spirometry in occupational health setting—2011 update. *J Occup Environ Med.* 2011;53:569–584.
3. Hankinson JL, Wagner GR. Medical screening using periodic spirometry for detection of chronic lung disease. *Occup Med.* 1993;8:353–361.
4. Wagner GR. *Screening and Surveillance of Workers Exposed to Mineral Dust*. Geneva, Switzerland: World Health Organization; 1996.
5. Harber P, Lockey JE. Pulmonary function testing in pulmonary prevention. *Occup Med State Art Rev.* 1991;6:69–79.
6. Hnizdo E, Berry A, Hakobyan A, Beeckman-Wagner LA, Catlett L. Worksite wellness program for respiratory disease prevention in heavy construction workers. *J Occup Environ Med.* 2011;53:274–281.
7. Hnizdo E, Glindmeyer HW, Petsonk EL. Workplace spirometry monitoring for respiratory disease prevention: a method review. *Int J Tuberc Lung Dis.* 2010;14:1–10.
8. Hnizdo E, Yu L, Freyder L, Lefante J, Glindmeyer HW. Precision of longitudinal lung function measurements—monitoring and interpretation. *Occup Environ Med.* 2005;62:695–701.
9. Sherrill DL, Lebowitz MD, Knudson RJ, Burrows B. Continuous longitudinal regression equations for pulmonary function measures. *Eur Respir J*. 1992;5:452–462.
10. Miller MR, Hankinson J, Brusasco V, et al. Standardization of spirometry. Standardization of lung function testing. *Eur Respir J.* 2005;26:319–338.
11. Enright PL, Johnson LJ, Connet JE, et al. Spirometry in the Lung Health Study: i. Methods and quality control. *Am Rev Respir Dis.* 1991;143:1215–1223.
12. Enright PL, Skloot GS, Cox-Ganser JM, Udasin IG, Herbert R. Quality of spirometry performed by 13,599 participants in the World Trade Center worker and volunteer medical screening program. *Respir Care.* 2010;55:303–309.
13. Fishwick D, Naylor S. COPD and the workplace. Is it really possible to detect early cases? Editorial. *Occup Med (Lond).* 2007;57:82–84.
14. Burgess JL, Nanson CJ, Bolstad-Johnson DM, et al. Adverse respiratory effects following overhaul in firefighters. *J Occup Environ Med.* 2001;43:467–473.
15. *Spirometry Longitudinal Data Analysis (SPIROLA) software*. http://www.cdc.gov/niosh/topics/spirometry/spirola.html. Accessed July 1, 2011.
16. Hnizdo E, Yan T, Enright P, et al. Spirometry Longitudinal Data Analysis software (SPIROLA) for Analysis of Spirometry Data in Workplace Prevention or COPD Treatment. *Open Med Inform J.* 2010;4:94–102.
17. NIOSH-Approved Spirometry Training Course. http://www.cdc.gov/niosh/topics/spirometry/training.html. Accessed July 1, 2011.
18. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general US population. *Am J Respir Crit Care Med.* 1999;159:179–187.
19. Beeckman-Wagner LF, Freeland DL, Shah Das M, Thomas KC. *Get Valid Spirometry Results Every Time*. Morgantown, WV: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health; 2011. Available at: http://www.cdc.gov/niosh/docs/2011-135/. Accessed August 26, 2011.
20. Miller MR, Crapo R, Hankinson J, et al. General considerations for lung function testing. *Eur Respir J.* 2005;26:153–161.
21. Hnizdo E, Yan T, Sircar K, Harber P, Fleming J, Glindmeyer HW. Limits of longitudinal decline for the interpretation of annual changes in FEV_1 in individuals. *Occup Environ Med.* 2007;64:701–707.
22. Wang ML, Vashia BH, Petsonk LE. Interpreting periodic lung function tests in individuals. The relationship between 1- to 5-year and long-term FEV_1 changes. *Chest.* 2006;130:493–499.
23. Wang M, Petsonk EL. Repeated measures of FEV_1 over six to twelve months: what change is abnormal? *J Occup Environ Med.* 2004;46:591–595.
24. Akpinar-Elci M, Travis WD, Lynch DA, Kreiss K. Bronchiolitis obliterans syndrome in popcorn production plant workers. *Eur Respir J*. 2004;24:298–302.

