AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 55:657–668 (2012)

Longitudinal Lung Function Declines Among California Flavoring Manufacturing Workers†

Kathleen Kreiss, MD,[1]* Kathleen B. Fedan, BS,[1] Muazzam Nasrullah, MD,[1,2] Thomas J. Kim, MD,[2,3] Barbara L. Materna, PhD,[3] Janice C. Prudhomme, DO,[3] and Paul L. Enright, MD[4]

Background *The California Department of Public Health received serial spirometry data for flavoring manufacturing workers at 20 companies at risk of bronchiolitis obliterans.*

Methods *We graded spirometry quality; identified individual workers with excessive decline in forced expiratory volume in 1 s (FEV_1) using relative longitudinal limits of decline based on 4% average within-person variability; and analyzed declines by occupational risk factors.*

Results *The quality of 1,696 spirometry tests from 724 workers varied by 17 providers, with poorer quality from commercial providers. Of 416 workers with at least two tests, 40 (9.6%) had abnormal FEV_1 decline. Of 289 workers with high quality spirometry, 21 (7.3%) had abnormal decline. Only one of the 21 had airways obstruction. Abnormal FEV_1 decline rates (per person-month) were greater among workers at companies using $\geq$800 lbs/year diacetyl than at companies using lesser amounts. Abnormal FEV_1 decline rates were greater at companies previously having four-person clusters of spirometric obstruction than at companies with no or only one worker with obstruction.*

Conclusions *Spirometric surveillance of flavoring workers can identify individual workers with an abnormal FEV_1 decline for preventive intervention, even when the FEV_1 itself remains within the normal range. Good quality spirometry and classification of abnormal with relative longitudinal limit of decline minimize misclassification of possible work-related health effects.* Am. J. Ind. Med. 55:657–668, 2012.

KEY WORDS: bronchiolitis obliterans; flavoring; diacetyl; spirometry; surveillance

[1]Division of Respiratory Disease Studies, National Institute for Occupational Safety and Health (NIOSH), Morgantown, West Virginia

[2]Epidemic Intelligence Service, Office of Workforce and Career Development, Centers for Disease Control and Prevention, Atlanta, Georgia

[3]Occupational Health Branch, Division of Environmental and Occupational Disease Control, California Department of Public Health (CDPH), Richmond, California

[4]College of Public Health, The University of Arizona, Tucson, Arizona

Disclosure Statement: None of the authors have conflicts of interest to report. This work was performed by State and Federal Government employees as part of their work; no non-governmental funding supported this work.

†This article was published online on 19 September 2011 and subsequently some errors based on the California Department of Public Health data included in the article were identified. The tables based on the faulty data have been corrected on 31 March 2012. The conclusions of the research have not been affected.

Muazzam Nasrullah present address is Office of Infectious Diseases (OID), National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention (NCHHSTP), Centers for Disease Control and Prevention, Atlanta, Georgia.

Thomas J. Kim present address is California Tribal Epidemiology Center, California Rural Indian Health Board, Sacramento, California.

Janice C. Prudhomme present address is Division of Occupational Safety and Health (Cal/OSHA), Department of Industrial Relations, Oakland, California.

*Correspondence to: Kathleen Kreiss, MD, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, 1095 Willowdale Road Mailstop H-2800, Morgantown, WV 26505. E-mail: kkreiss@cdc.gov

Accepted 30 August 2011

DOI 10.1002/ajim.21013. Published online 19 September 2011 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

Inhaling diacetyl (2,3-butanedione) and artificial butter flavoring powders and vapors can cause fixed obstructive lung disease and bronchiolitis obliterans in flavoring-exposed workers [Kreiss, 2007]. Drafting of federal regulations is underway to prevent occupational lung disease in workers in flavoring manufacturing, microwave popcorn, and other food production industries. Even after an occupational exposure limit is established for diacetyl, medical surveillance will remain important for these workers because substitutes that are themselves potentially hazardous are now being used in place of diacetyl [Day et al., 2011]. The quantitative toxicity of these substitutes is still unknown, but 2,3-pentanedione, another alpha-diketone, has epithelial toxicity comparable to diacetyl in animal models [Hubbs et al., 2010; Morgan et al., 2011]. Thus, longitudinal testing of exposed workers for evidence of excessive decline in forced expiratory volume in 1 s (FEV_1) may be the only way to assure early identification of possible adverse health effects and to motivate exposure control for both affected workers and co-workers at risk from exposure to potentially hazardous flavoring chemicals.

The major public health experience in monitoring exposed workers in the flavoring industry over time has been in California. Starting in 2006, the California Department of Public Health (CDPH) asked flavoring manufacturing companies to have their medical providers submit personally identifiable spirometry and questionnaire data at 3- to 6-month intervals, as part of a joint program with the California Division of Occupational Safety and Health (Cal/OSHA) to protect the health of flavoring workers [Kim et al., 2010]. The primary intent of this public health surveillance was to follow back with medical providers and companies when abnormalities in spirometry suggested the need for further clinical evaluation of individuals and preventive measures in the factory. A second goal was to establish the extent of flavoring-related morbidity in California and potential risk factors that could result in prevention of disease. Initial cross-sectional analyses of submitted data identified 18 workers with spirometric obstruction, at least 12 of whom had probable occupational fixed airways obstruction. Compared to the general population, flavoring workers were 2.7 times more likely to have severe airways obstruction. Risk factors for obstruction included younger age, Hispanic ethnicity, liquid and powder production work, greater company diacetyl usage, and having a co-worker with obstruction [Kim et al., 2010].

This report presents analyses of submitted data with the following objectives: (1) to assess the quality of spirometry by company and provider; (2) to determine the effect of spirometry quality on worker numbers judged to have excessive FEV_1 decline; (3) to quantify the longitudinal change in FEV_1 in terms of the mean rate of decline and the rate of excessive declines by work-related risk factors; and (4) to distill implications of the California industry-wide spirometry surveillance for future monitoring efforts for workers at risk of rapidly progressive occupational lung diseases.

METHODS

Setting and Data Sources

CDPH disseminated medical provider guidelines, including a worker questionnaire with validated questions [CDPH, 2007], standards for spirometry testing, instructions for spirometry reporting [Miller et al., 2005], and suggested clinical evaluation for those found to have spirometric abnormalities. Although CDPH recommended that surveillance be supervised by a physician who is board-certified in occupational or pulmonary medicine and promoted the use of university-affiliated specialty clinics, flavor manufacturers chose their medical service providers. CDPH has authority to receive personally identifiable information subject to the Health Insurance Portability and Accountability Act, which permits medical providers to disclose protected health information, without the written authorization of the patient to whom the information pertains, to state health departments or other public health authorities that are authorized by law to collect or receive such information for the purpose of preventing or controlling disease, injury, or disability [U.S. Code of Federal Regulations, 2002]. The institutional review board of the National Institute for Occupational Safety and Health (NIOSH) ruled that this public health surveillance effort was not research (that would have required informed consent from workers); similar judgments were made at the Centers for Disease Control's Office of Workforce and Career Development and at CDPH regarding its non-research status.

A total of 17 medical providers to 20 flavoring manufacturing companies submitted hardcopy questionnaire and spirometry results collected from 4/13/2004 to 5/8/2009, in many cases years after flavoring workers had been interviewed and tested. All workers could be linked to an employer which allowed us to examine company-related potential risk factors such as quantity of diacetyl used in 2006 (at the inception of the Cal/OSHA prevention program) and clustering of obstructive lung disease. Previously extracted questionnaire data through March 2008 [Kim et al., 2010] were available for 326 (78.4%) of the 416 workers with serial tests. With this information, we were able to assign 320 workers by their highest exposure task into current production, production support, or minimally exposed job categories. Likewise, we were able to assign 205 of these 416 workers as currently working

exclusively in production, production support, or minimally exposed job categories.

Spirometry Quality Scoring

Per medical surveillance guidelines [CDPH, 2007], spirometers were to be configured to print the results from the best three expiratory maneuvers. The submitted spirometry reports were produced by a variety of spirometer types; therefore reports did not share a common format or contain the same information. All trials (i.e., expiratory maneuvers) were reviewed for acceptability by our spirometry expert (PE). From each submitted hardcopy spirometry test report, we recorded the number of trials displayed, the number of trials judged acceptable on the report, whether the report provided information on each trial's extrapolated volume (used to correct the FEV_1 for slow starts), whether the report indicated if the test demonstrated repeatability of the best trials, and the highest two values of both FEV_1 and forced vital capacity (FVC) from maneuvers judged acceptable by PE and the spirometer (when available). We calculated the absolute variability (in milliliters) of FEV_1 and FVC from these recorded observations to assess repeatability and assign quality scoring. For FEV_1, we assigned quality grade A or B if the highest two values were within 100 ml or within 101–150 ml, respectively, of each other, and we defined these two grades as evidence of a good quality test for FEV_1 because they met American Thoracic Society repeatability guidelines [Miller et al., 2005]. We assigned grade C if the values were within 151–250 ml of each other. We assigned grade D if only one trial was judged as acceptable or if the values of two acceptable trials differed by more than 250 ml. For submitted spirometry test reports that displayed and/or reported only a single maneuver, we could not independently determine quality based on repeatability. For these tests, we assigned grade C− if the report indicated that there were at least two acceptable trials and that these demonstrated repeatability (although the level of repeatability was unknown). We assigned grade D− if the report indicated that only one trial was acceptable. We assigned grade F if the report indicated or PE determined that the test had no acceptable trials. We assigned no quality grades to tests for which reports displayed results of only one trial *and* the reports did not contain the information needed to determine acceptability or repeatability. FVC test results were graded for quality utilizing the same criteria as FEV_1.

Interpretation of Spirometry Test Results

We defined spirometric restriction as FVC below the lower limit of normal (LLN) with a normal FEV_1/FVC ratio, as determined by NHANES III reference equations [Hankinson et al., 1999]. We defined spirometric obstruction as FEV_1/FVC *and* FEV_1 below the LLN. We included workers with both spirometric obstruction and spirometric restriction (i.e., a ''mixed'' pattern) in the obstructed group, because the low FVC in 90% of such cases among general patient populations is due to hyperinflation and not to a low total lung capacity [Dykstra et al., 1999].

Calculation of Longitudinal FEV_1 Change and Determination of Abnormal FEV_1 Decline

For the longitudinal analyses, we excluded workers who did not have at least two spirometry test dates 3 or more months apart and examined longitudinal change in lung function using two approaches, one continuous and one categorical. For each worker we calculated the following continuous variables: the absolute change in FEV_1 (last measurement − first measurement), the relative change in FEV_1 (absolute change in FEV_1 divided by the first measurement and then multiplied by 100%), and the slope of FEV_1 decline (estimated from simple linear regression and using all measurements for each individual) expressed as both the change over time (units = ml/year) and as the percent change over time (units = %/year).

In the second approach, we classified each worker as having a normal or abnormal FEV_1 decline categorically using SPIROLA [Hnizdo et al., 2007; NIOSH, 2010], a free-ware program developed by NIOSH to monitor and evaluate lung function longitudinally. SPIROLA was used to (1) calculate the average pair-wise within-person variation of the data and (2) to use this estimate to determine and apply the relative longitudinal limit of decline to categorize abnormal FEV_1 decline. The good quality data had an average within-person variation of 4%, with which we made the comparison for objective 2 (to determine the effect of spirometry quality on worker numbers judged to have excessive FEV_1 decline). All spirometry data had an average within-person variation of 5% with a corresponding higher threshold for categorical classification of having abnormal FEV_1 decline.

Statistical Analysis

We prepared distributions of data quality by provider, type of provider (i.e., academic or government vs. commercial), and company. We examined abnormal FEV_1 declines and group mean FEV_1 changes by potential occupational risk factors shown to be associated with prevalence of spirometric obstruction in previous cross-sectional analyses of this population. Occupational risk factors were company use of $\geq$800 lbs of diacetyl in 2006; companies with four-person clusters of obstructive

abnormalities (having a co-worker with obstruction) in cross-sectional analyses; current production job; and currently performing any production task [Kim et al., 2010]. Using the spirometry records submitted to CDPH for each company, we adjusted for duration of company medical surveillance and workforce size by calculating frequency of abnormal FEV_1 decline using person-months of surveillance by company. One set of analyses included all data regardless of quality, and a second set of analyses included only tests with good quality FEV_1 data. We used SAS/STAT® software [SAS, 2004] to conduct all statistical analyses. Chi-square was used for categorical data and Student's *t*-test for continuous data. Differences between rates of abnormal FEV_1 decline were tested using a normal approximation to the Poisson distribution. We considered a two-tailed *P*-value of ≤ 0.05 to be statistically significant and higher *P*-values up to 0.1 as marginally significant.

RESULTS

Worker Characteristics

Surveillance data submitted to CDPH included 585 (80.8%) male and 139 (19.2%) female flavoring workers, with a median age at their first reported test session of 37.0 years (range: 17–69 years). There were 680 (93.9%) workers with known race and ethnicity: 387 (53.5%) Hispanic, 150 (20.7%) non-Hispanic White, 31 (4.3%) non-Hispanic Black, 88 (12.2%) Asian, and 24 (3.3%) of other race or ethnicity. Workers for whom more than one spirometry test result was submitted had similar gender, race, and ethnicity distributions, but they were significantly older than workers tested only once (average age of 38.8 years vs. 35.0 years, $P < 0.0001$). Cigarette smoking information was available for 77% of the workers with serial spirometry measurements: 17.8% were current smokers and 17.5% were former smokers. Participants working in companies using $\geq$800 lbs/year diacetyl did not differ significantly from participants working in companies using less diacetyl in percentage of ever-smokers (38.1% vs. 31.3%, $P = 0.24$), but the workers with missing smoking information were not evenly distributed between the high and low diacetyl-using companies. Only 35 (15.6%) of those missing smoking information worked in companies using $\geq$800 lbs/year diacetyl, whereas 61 (31.8%) worked in companies using $<$800 lbs/year diacetyl.

Of 724 workers for whom spirometry data were submitted, we could not confidently interpret the results from 47 because information on race/ethnicity was missing. Of the remaining 677 workers, 156 (23.0%) had abnormal spirometry during one or more tests, including 33 (4.9%) with airways obstruction. Of the 416 workers with serial tests (i.e., more than one), 411 had interpretable tests, and 19 (4.6%) had at least one test indicating airways obstruction. We previously reported 18 workers with airways obstruction detected during cross-sectional analyses of workers at 16 of the 20 companies [Kim et al., 2010]. Nine of these workers (including all five with severe obstruction) had no additional spirometry tests submitted as part of workplace screening.

Company Data

Companies varied with respect to the period of time covered by submitted spirometry data (Table I). Testing was not always submitted for every year after testing was initiated, despite guidance that testing was to be performed at least every 6 months and as often as every 3 months for companies with recognized cases of work-related fixed obstruction. Fifty-five percent of companies had more than one medical provider. Some companies changed providers during the surveillance period in response to feedback from CDPH staff that poor quality tests had been submitted.

Medical Provider Data

The 17 medical providers submitted spirometry tests on 724 workers at the 20 flavoring companies (Table II). Six providers contributed 74% of all tests. Serial tests were available for 416 workers: 193 (46.4%) with 2 tests, 149 (35.8%) with 3–4 tests, 44 (10.6%) with 5–6 tests, and 30 (7.2%) with 7–8 tests. Intervals between first and last tests for each worker ranged from 3 to 57 months, with a mean of 20 months and a median of 16 months. Among workers with serial tests, about two-thirds (65.9%) had no more than 2 years of serial testing submitted, and only 3.8% had more than 4 years. Of 416 workers with serial tests, 215 (51.7%) were tested by two providers, and 30 (7.2%) were tested by three providers.

Spirometry Quality by Provider and Company

The medical providers used 16 different models of spirometers; the number of tests performed on each spirometer model ranged from 8 to 309. Almost all of the spirometers used by the medical providers can be configured to store and print the numeric and graphic results of the three best maneuvers within a test session, per CDPH [2007] and American Thoracic Society guidelines [Miller et al., 2005], but in many cases, only results from the single "best" maneuver were submitted.

Overall, 71% of the tests received good quality grades (i.e., grade A or B) for FEV_1 (Table II). There were 109

TABLE I. Summary of 1,696 Spirometry Tests for 724 Workers at the 20 Flavoring Companies, of Whom 416 Had Serial Tests

Company	Providers[a]	Tests submitted (n)	Tests with good FEV_1 quality (%)	Workers tested[b] (n)	Workers with serial tests[b] (n)	Years tests performed
Diacetyl use ≥800 lb/year						
1	C,B	256	97.3	73	46	2006–2009
2	A,I	191	76.4	130	61	2006–2007
3	A	188	93.6	81	49	2005–2007
4	A,F	184	23.9	67	35	2005–2009
5	C,M	71	87.3	42	17	2006–2009
6	J,D	30	36.7	21	9	2007–2008
7	L	20	35.0	13	7	2006
Diacetyl use <800 lb/year						
8	A,D,B	145	61.4	49	40	2006, 2008, 2009
9	A,H	142	40.8	51	34	2005–2008
10	G	114	82.5	30	19	2004–2009
11	A,B	86	95.3	32	24	2006–2009
12	E,J	79	82.3	41	23	2006–2008
13	K	49	53.1	29	20	2007–2009
14	B	40	95.0	21	10	2006–2008
15	A,N,P	38	60.5	12	6	2006–2009
16	A,O	33	30.3	14	9	2005–2009
17	B	16	93.8	11	3	2006–2007
18	B	6	83.3	3	3	2006–2007
19	B	5	80.0	1	1	2006–2007
20	Q	3	66.7	3	0	2009
All 20 companies						
1–20	A–Q	1,696	71.1	724	416	2004–2009

[a]Medical provider codes are consistent with those used in Table II and are listed in the chronological order flavoring employees were tested at each company.
[b]Of 416 workers with multiple tests, 408 (98.1%) were tested while working for the same flavoring company throughout the surveillance period. In this table, the 8 workers that worked at more than one company were assigned to the company providing their most recent test.

(6%) tests that were either not gradable or were graded F for either FEV_1 or FVC. No lung function measurements were usable for one test due to a malfunction in the spirometer; no FVC values were usable for five additional tests because all available trials contained extra breaths, although the FEV_1 values for these tests were usable as the extra breaths occurred after the first second of each maneuver. Quality distributions for the 1,387 tests on the 416 workers with two or more spirometry tests were similar to those for the 308 workers with only one test (data not shown).

Quality of submitted spirometry tests differed by provider (Table II). Good quality FEV_1 tests ranged from 0% to 99.1% of all tests by providers submitting at least 10 tests; ungradable or F quality FEV_1 tests ranged from 0% to 68.5% for these providers. Eight providers submitted reports with no information on acceptability of trials, and three other providers submitted reports with acceptability information for <15% of tests. Similarly, 11 providers submitted little or no information on repeatability. The three non-profit (academic and government) medical providers submitting more than 100 tests each produced good quality FEV_1 tests for at least 94% of submitted tests. Of the three commercial providers submitting more than 100 tests each, one had 82.5% good quality tests, one had 22.6% good quality tests, and one did not configure their spirometers to print the results of the three best tests so repeatability and acceptability could not be verified (0.7% documented good quality).

As expected, companies varied in the quality of spirometry tests submitted by their medical providers (Table I). Good quality FEV_1 data comprised at least 80% of tests submitted for half the companies. However, for two of the six companies represented by the largest numbers of workers with serial FEV_1 data, accounting for 24% of all person-months of follow-up in our analyses, rates of good quality tests were substantially lower—28.6% and 47.1% of tests (data not shown).

TABLE II. Summary of Spirometry Tests Submitted by the 17 Participating Medical Providers

Provider	Tests submitted (n)	FEV_1 Quality Grade (% of tests submitted by provider)							Number of spirometer models	Trial acceptability documented[b] (%)	Trial repeatability documented[c] (%)	Workers with serial tests[d] (n)
		A/B	C	C−	D	D−	F	NG[a]				
Academic or government provider												
A	459	94.1	4.1		1.3	0.2		0.2	2	0.2	0.4	79
B	322	95.0	2.2	0.6	1.6		0.3	0.3	3	83.2	91.0	72
C	112	99.1			0.9				1	0	88.4	34
D	89	30.3	1.1				2.2	66.3	2	0	0	35
E	20	70.0	10.0		15.0		5.0		1	0	10.0	0
Commercial provider												
F	140	0.7		85.7		12.9		0.7	2	99.3	99.3	33
G	114	82.5	2.6		3.5	6.1	5.3		1	13.2	7.0	19
H	106	22.6	2.8		6.6	66.0		1.9	3	67.9	3.8	27
I	95	61.1	10.5		14.7	1.1	8.4	4.2	2	2.1	1.1	0
J	78	79.5	11.5		3.8	1.3	3.8		1	98.7	98.7	22
K	49	53.1	30.6		10.2		6.1		1	0	0	20
L	20	35.0	20.0		40.0		5.0		1	0	0	7
M	37	75.7	8.1		5.4		10.8		2	81.1	81.1	13
N	24	45.8	25.0		20.8			8.3	1	0	0	5
O	23	0.0		47.8		52.2			2	100	65.2	8
P	5	60.0	20.0		20.0				1	0	0	0
Q	3	66.7	33.3						1	0	0	0
All17 providers												
A–Q	1,696	71.1	5.0	7.8	3.8	6.5	1.7	4.1	1–3	37.0	39.5	416

[a]Test not gradable due to insufficient information on spirometry report.
[b]Spirometry report indicated the number of acceptable trials in the test.
[c]Spirometry report indicated whether the test met repeatability criteria.
[d]Number of workers tested two or more times by the same medical provider.

Abnormal FEV_1 Declines

Basing determinations on an analysis of spirometry data that did not consider test quality, 40 (9.6% of 416) workers were identified with abnormal FEV_1 decline (Table III), ranging from 12.7% (−200 ml) over 8 months to 49.3% (−1,700 ml) over 8 months. These 40 included 4 workers with obstructive abnormality (including one with a mixed abnormality), 13 with restrictive abnormality, 22 with normal spirometry, and 1 without race information. Workers with abnormal FEV_1 declines clustered by company: one company had seven such workers; one company had five; and two companies had four. Of the four companies with four-person clusters of obstructive spirometry in the cross-sectional study, three had at least three cases of abnormal FEV_1 decline. Companies using $\geq$800 lbs diacetyl in 2006 had a statistically significant excess of abnormal FEV_1 decline compared to companies using less diacetyl (7.3 vs. 3.0 per 1,000 person-months; $P = 0.01$) (Table IV). The mean follow-up interval for workers in these companies was 15.9 months, compared to 24.3 months in companies using less diacetyl ($P < 0.0001$). The proportion of workers having more than two measurements differed between the two subgroups (42.9% of the workers in companies using $\geq$800 lbs vs. 66.1% of the workers in companies using less, $P < 0.0001$). The seven companies using $\geq$800 lbs/year of diacetyl accounted for 26 of the 40 workers with abnormal FEV_1 decline; 11.6% (26 of 224) of their workers had abnormal FEV_1 decline compared to 7.3% (14 of 192) of workers at companies using less diacetyl. Among the subset of workers for which we had information on current work tasks abstracted from questionnaires, working exclusively in production or working in any production task carried about twice the rate of abnormal FEV_1 decline than working exclusively in production support or in minimally exposed jobs, but these differences were not statistically significant.

Basing determinations on good quality data only, 21 (7.3% of 289) workers were identified with abnormal

TABLE III. Abnormal FEV_1 Declines by Flavoring Company and Spirometry Quality

	Analysis of all tests regardless of quality				Analysis restricted to tests of good quality		
		Abnormal FEV_1 decline				Abnormal FEV_1 decline	
Company	Person-months of follow-up[a]	n[a]	Rate per 1,000 person-months	Number of workers with obstruction in the initial cross-sectional analysis[b]	Person-months of follow-up[a]	n[a]	Rate per 1,000 person-months
Diacetyl use ≥800 lb/year							
1	824	3	3.6	4	817	3	3.7
2	604	7	11.6	4	356	1	2.8
3	529	5	9.5	4	495	6	12.1
4	1019	4	3.9	0	145	0	0
5	349	2	5.7	4	278	3	10.8
6	190	2	10.5	0	26	0	0
7	56	3	53.6	NA	16	1	62.5
1–7	3,570	26	7.3	16	2,132	14	6.6
Diacetyl use <800 lb/year							
8	1,119	1	0.9	0	876	1	1.1
9	957	1	1.0	1	426	2	4.7
10	636	1	1.6	NA	540	0	0
11	662	1	1.5	0	647	1	1.5
12	320	3	9.4	NA	254	1	3.9
13	253	4	15.8	1	78	0	0
14	167	0	0	0	164	0	0
15	173	2	11.6	0	104	2	19.3
16	332	1	3.0	0	33	0	0
17	20	0	0	0	20	0	0
18	24	0	0	0	16	0	0
19[c]	7	0	0	0	—	—	—
20	0	—	—	NA	0	—	—
8–20	4,670	14	3.0	2	3,157	7	2.2
All 20 companies							
1–20	8,240	40	4.9	18	5,290	21	4.0

NA, companies not included in the cross-sectional analysis [Kim et al., 2010].
[a]Workers with tenure at more than one company are included in the company that submitted their most recent test.
[b]Kim et al. [2010].
[c]Company 19 submitted no serial tests with good quality.

declines (Table III), ranging from 9.8% (−440 ml) over 6 months to 49.3% (−1,700 ml) over 8 months. These 21 included 1 with obstruction (a worker who worked exclusively at production tasks and developed obstruction while under surveillance), 5 with restriction, 14 with normal spirometry, and 1 without race information. The worker with obstruction had an FEV_1 decline of 23.9% (−980 ml) over 25 months. Workers with abnormal FEV_1 declines clustered by company: One company had six cases; two had three cases; and two had two cases. Companies using ≥800 lbs diacetyl in 2006 had a significant excess of abnormal FEV_1 decline compared to companies using less diacetyl (6.6 vs. 2.2 per 1,000 person-months; $P = 0.03$) (Table IV). The mean follow-up interval for workers in companies using ≥800 lbs diacetyl was 13.8 months compared to 23.6 months in the companies using less diacetyl ($P < 0.0001$). The proportion of workers having more than two measurements differed between the two subgroups (49.0% of the workers in companies using ≥800 lbs vs. 81.3% of the workers in companies using less, $P < 0.0001$). The seven companies using ≥800 lbs of diacetyl in 2006 accounted for 14 (66.7%) of the 21 workers with abnormal decline; 9.0% (14 of 155) of their workers had abnormal decline compared to 5.2% (7 of 134) at companies using less diacetyl.

TABLE IV. Abnormal FEV_1 Declines by Exposure Risk Factors and Spirometry Quality

	Analysis of all tests regardless of quality					Analysis restricted to tests of good quality				
				Abnormal FEV_1 decline					Abnormal FEV_1 decline	
Exposure risk factor	Workers with serial tests (n)	Mean FEV_1 decline (ml/year)	Person-months of follow-up	n	Rate per 1,000 person-months	Workers with serial tests (n)	Mean FEV_1 decline (ml/year)	Person-months of follow-up	n	Rate per 1,000 person-months
High diacetyl use[a]	416					289				
≥800 lbs/year	224	86.4	3,570	26	7.3[d]	155	114.3[e]	2,132	14	6.6[d]
<800 lbs/year	192	53.6	4,670	14	3.0	134	50.4	3,157	7	2.2
Obstruction clusters[a]	367					251				
Four cases	173	63.3	2,305	17	7.4[d]	142	96.0	1,946	13	6.7[d]
<Four cases	194	47.5	4,923	16	3.2	109	50.8	2,534	6	2.4
Current work tasks—hierarchal[b]	320					234				
Production	225	53.7	4,550	23	5.1	165	67.5	3,046	14	4.6
Production support	72	44.2	1,655	4	2.4	53	89.6	969	2	2.1
Minimally exposed	23	67.9	384	1	2.6	16	63.6	236	1	4.2
Current work tasks—exclusive[c]	205					142				
Production	120	50.5	2,105	12	5.7	82	89.2	1,366	8	5.9
Production support	62	60.5	1,370	3	2.2	44	109.6	791	2	2.5
Minimally exposed	23	67.9	384	1	2.6	16	63.6	236	1	4.2

[a]Workers with tenure at more than one company are included in the company that submitted their most recent test.
[b]Workers assigned to category based on work task associated with highest exposure [Kim et al., 2010].
[c]Workers assigned to category if all current work tasks were in that category [Kim et al., 2010].
[d]Statistically different at $P \leq 0.05$.
[e]Marginally statistically different at $0.05 < P \leq 0.1$.

Four of the 21 workers with abnormal FEV_1 decline based on good quality data were not among the 40 workers with abnormal FEV_1 decline based on all data. (Their FEV_1 declines based on all data were not sufficiently negative to meet criteria for abnormal FEV_1 decline, even though the good quality data subset showed that an abnormal decline occurred during their monitoring period.) Two of these four were in groups determined in our previous cross-sectional analyses [Kim et al., 2010] to be at "high-risk" for obstruction (i.e., company use of ≥800 lbs/year diacetyl; and company with a four-person cluster of obstructive abnormalities). Thus, a total of 44 (10.6%) of the 416 workers with at least two spirometry tests were identified as having abnormal FEV_1 decline using either all the spirometry data or the subset of good quality spirometry.

Using the within-person variability of 5% for all quality spirometry resulted in 28 (6.7% of 416) persons having excessive decline below the relative longitudinal limit of decline, in comparison to the 40 (9.6%) persons identified using the 4% within-person variability.

Absolute and Relative FEV_1 Change

The 416 workers with two or more tests experienced an overall mean absolute FEV_1 change from first to last test of −98.2 ml (range: −1,810 to +830 ml), corresponding to a relative percent FEV_1 change of −2.4% (range: −49.3% to +35.8%). When normalized to annual FEV_1 loss, the mean absolute change over the submitted spirometry period was −71.3 ml/year (range: −2534.4 to +2435.0 ml/year). The 19 workers with at least one obstructed test (4.6%) had an average FEV_1 slope of −220.6 ml/year versus −56.0 ml/year for those workers without an obstructed test ($P = 0.23$).

Restricting analysis to good quality data, the 289 workers with at least two good quality tests experienced an overall mean absolute FEV_1 change from first to last test of −113.6 ml (range: −1,810 to +560 ml), corresponding to a relative percent FEV_1 change of −2.9% (range: −49.3% to +31.4%). When normalized to annual FEV_1 loss, the absolute change over the submitted spirometry period was −84.7 ml/year (−2534.4 to +926.8 ml/year). The 11 workers with at least one obstructed test (3.8%) had an average FEV_1 slope of −97.5 ml/year versus −74.3 ml/year for those workers without an obstructed test ($P = 0.76$). Consistent with the excess of abnormal FEV_1 declines among workers at companies using ≥800 lbs/year diacetyl, the mean FEV_1 change for companies using ≥800 lbs/year of diacetyl was −114.3 ml/year compared to −50.4 ml/year in companies using less diacetyl ($P = 0.06$).

DISCUSSION

Voluntary industry-wide reporting of serial spirometry for 416 flavoring-exposed workers to the California Department of Public Health resulted in identification of 40 (9.6%) workers with abnormal declines in FEV_1. Whether limiting analyses to good quality spirometry pairs (7.3% of 289 workers with abnormal FEV_1 decline) or including all quality spirometry (9.6% of 416 workers with abnormal FEV_1 decline), these percentages exceed the 5% criterion on which the upper limits of normal FEV_1 decline are based. However, what does examination of abnormal FEV_1 decline potentially contribute to the health of these flavoring workers? The answer to this question depends on three factors: (1) Are additional workers identified by serial testing as having abnormal FEV_1 declines who do not have obstructive abnormalities on a single test? (2) Are workers with abnormal FEV_1 declines likely to have an occupational cause? (3) Are doctors, companies, and employees likely to respond to abnormal FEV_1 declines with preventive measures to prevent progression to impairment?

Nine percent of workers with abnormal FEV_1 declines in serial spirometry data had obstructive abnormalities during the period covered by submitted spirometry. The person with abnormal FEV_1 and abnormal FEV_1 decline using A and B quality spirometry developed obstruction during surveillance on the occasion of his second test. Thus, the workers identified in serial spirometry testing as having abnormal FEV_1 declines are additional workers who may have experienced lung function loss within the normal range of FEV_1 as a result of flavoring exposures. The minimal overlap between the 18 workers found to have obstruction in the cross-sectional study and the 21 workers found to have abnormal FEV_1 decline with good quality spirometry in the current study suggests that the obstructed workers may have left the industry, been reassigned to work that was not considered to merit surveillance testing, or had no further abnormal fall in FEV_1. Follow-up of workers with abnormal declines in FEV_1 is intended to limit further flavoring exposure for those with a possible occupational explanation, in order to prevent them from developing irreversible fixed obstruction from their workplace exposure.

The evidence for occupational cause in a portion of those we identified with abnormal FEV_1 declines is threefold. First, the prevalence of abnormal FEV_1 decline is high in this worker population, up to 10.6%, depending on the quality subgroups considered. It is unlikely that workers with a predisposition to developing airways obstruction are preferentially employed in industries with flavoring exposures. Second, a causal temporal sequence exists for flavoring exposures preceding the occupational health effect, as measured in serial FEV_1 decrements. Third, these surveillance results demonstrate that a priori hypotheses regarding occupational risk factors for abnormal FEV_1 decline were affirmed. Companies with greater potential for diacetyl exposures as a consequence of annual poundage used in 2006 had higher prevalence of abnormal FEV_1 declines in their workers compared to companies using less diacetyl. Workers with good quality spirometry who worked in companies using more diacetyl tended to have double the average annualized FEV_1 decline of workers in companies using lesser amounts of diacetyl. The companies with four-person clusters of workers with obstruction in cross-sectional analyses [Kim et al., 2010] (a subset of the companies using $\geq$800 lbs/year diacetyl) had significantly higher rates of abnormal FEV_1 decline. So although not all workers with abnormal FEV_1 declines may have an occupational cause related to flavoring exposures, the occupational risk factors are those associated with obstructive abnormalities demonstrated in the cross-sectional analyses [Kim et al., 2010].

For serial spirometry to achieve its potential role in preventing the development of irreversible fixed airways obstruction due to flavoring exposures requires action on the part of all those participating in medical surveillance. Action on the part of physicians in response to abnormal FEV_1 declines has been inapparent to date. Some physicians identifying mild and even moderate obstruction among California flavoring workers did not report diagnostic evaluations of these patients to the California Department of Public Health [Kim et al., 2010], despite repeated prompting by health department physicians. Some of these patients with airways obstruction did not even have submitted evidence of post-bronchodilator spirometry to differentiate fixed obstruction from more common lung diseases such as asthma. The majority of mild and moderate cases of obstruction had no chest symptoms, perhaps resulting in reluctance of workers and their physicians to seek further evaluation. The incomplete clinical information submitted by medical providers for some flavoring companies is not reassuring regarding whether these workers with airways obstruction were evaluated clinically or counseled regarding protective measures.

Companies volunteered to participate in medical surveillance in exchange for avoiding Cal/OSHA compliance visits. In early 2009, Cal/OSHA restated in writing its intention to enforce the agreement, at which point half of the companies had their providers submit some additional surveillance data, which have been incorporated into this article. Thus, longitudinal analyses of serial data by public health authorities as the tests were performed were impractical, and little evidence existed that medical providers were evaluating serial measurements to detect abnormal declines. Indeed, with the frequent changes in medical providers, serial data were not consistently available to providers to evaluate for interval FEV_1 declines. Lack of

comparability among spirometers within and between providers may have also played a role in interpreting longitudinal changes [Townsend and Occupational and Environmental Lung Disorders Committee, 2011]. Lastly, economic and workplace pressures on flavoring workers, about half of whom are Hispanic in California, may have contributed to difficulty in implementing preventive measures. One provider only notified employers of workers with severe abnormalities in FEV_1 because of concern for protecting continued employment of those with lesser degrees of impairment. All these factors diminish the likelihood of preventive action in response to individual abnormal FEV_1 declines within the normal FEV_1 range in the current setting without regulation dictating surveillance procedures and ensuring spirometry quality.

The results of our analysis of serial spirometry data document challenges for most medical providers in providing good quality spirometry. Although some university-affiliated medical providers and government teams provided top quality spirometry, some non-profit providers and almost all commercial providers provided either dismal quality spirometry or spirometry whose quality could not be independently assessed. Quality grading guidelines were set by the American Thoracic Society/European Respiratory Society [Miller et al., 2005] with the knowledge that 90% of adult patients can achieve A or B quality spirometry with a competent technician [Enright et al., 2004]. A more recent evaluation of technicians performing spirometry for World Trade Center rescue workers argues that at least 80% of spirometry test sessions should yield good quality results as a criterion of adequate technician performance [Enright et al., 2010]. In only one of 12 commercial spirometry providers to California flavoring companies could we document this level of spirometry quality. While the importance of spirometry in surveillance is undoubtedly recognized by most clinicians, the level of attention to spirometry quality clearly varied, consistent with most providers having inadequate quality assurance programs. Physicians demonstrated poor control practices by reviewing and interpreting poor quality tests and not requiring repeat testing. Perhaps some physicians always accepted the computer interpretations provided by the spirometer.

While spirometry quality may be poor due to lack of proper spirometry technologist training, of continued education and practice, of patient cooperation, and the spirometer itself, the supervising medical provider as the end-user of these data serves as the final gate-keeper of spirometry quality. The poor quality data submitted by commercial and some academic providers may indicate deficiency in spirometry education and interpretation skills [Leuppi et al., 2010]. Indeed, one test submitted was interpreted as normal with an FEV_1 of 7.21 L, a physiologically impossible result for a man of normal stature. A day or two of spirometry training is insufficient to guarantee adequate performance by either physician or technologist [Borg et al., 2010]. A performance-based certificate of quality for technologists has been instituted [AARC, 2010] and may be useful in increasing the likelihood of a quality product. Companies contracting for medical surveillance examinations might consider fulfillment of quality criteria as a condition for payment. In the experience of the California flavoring worker surveillance, nine of the 17 (53%) providers would not have merited payment under such contractual language which would require at least reconfiguring their spirometer systems to print the numeric and graphic results from the three best spirometry maneuvers so that quality could be independently determined.

The number of workers identified with abnormal FEV_1 declines was influenced by spirometry quality, with use of only good quality FEV_1s resulting in identification of about half (21) the number of workers in comparison to using all spirometry data with the same evaluation criterion reflecting an average 4% within-person variability. When 5% within-person variability was used in the SPIROLA program to calculate the relative longitudinal limit of decline, the number of persons identified as having abnormal FEV_1 declines in all quality data was higher than the number using the 4% criterion in the high quality subset, but the rate of abnormality was lower among the 416 with serial measurements. Including poor and unverifiable quality test results may have resulted in some false positives. However, it is reassuring that results regarding work-related risk factors were comparable in analyses using all quality and good quality data, and it is possible that some unverifiable quality tests may have actually been good quality. Misclassification of health outcome because of poor quality spirometry results in unnecessary referral of workers for further medical evaluation. Money wasted on poor quality medical surveillance impacts employers and may impact workers who may be inappropriately restricted from further flavoring exposure.

The optimal frequency of periodic spirometry testing to detect airways obstruction in flavoring workers depends on several factors, including the pretest probability of disease, how rapidly the disease progresses, and weighing the consequences of false-positive and false-negative interpretations of test results [Harber et al., 2009; Hnizdo et al., 2010]. One worker in this study lost half of his lung function in 8 months, and an incident case of obstruction lost 1 L of FEV_1 in 4.5 months. When a disease such as bronchiolitis obliterans can progress very rapidly, more frequent than annual testing of high-risk workers is advisable. In weighing sensitivity versus specificity [Wang and Petsonk, 2004; Wang et al., 2006; Townsend and Occupational and Environmental Lung Disorders Committee, 2011], risk factors such as the occurrence of a case of flavoring-related lung disease, large diacetyl or diacetyl

substitute use, production work, and results of previous spirometry, support the choice of shorter testing intervals. Spirometry quality, reflected in low within-person variation, is the critical underpinning of being able to confidently identify excessive FEV_1 decline during the short periods of follow-up that are pertinent to development of flavoring-related lung disease [Hnizdo et al., 2010].

The limitations of this surveillance system were many. The longitudinal results are unlikely to be representative, even of California flavoring workers. Companies had their medical providers submit surveillance data for differing durations of surveillance, which resulted in varying contributions to person-months of longitudinal follow-up (apart from number of workers). The average duration of spirometry monitoring from companies using $\geq$800 lbs/year of diacetyl was shorter than from companies using less diacetyl. These differences may have affected the precision of our estimates of FEV_1 decline and thus our comparison of abnormal FEV_1 decline rates. In analyses of good quality data, our results were weighted towards workers in companies whose providers submitted good quality data; workers at companies contracting with commercial spirometry providers are underrepresented in our serial spirometry analyses using good quality data. Because providers generally did not use the CDPH questionnaire, we did not have complete data on worker job title, tasks, or even smoking history; we lacked questionnaire data on 22% of those with serial testing. Some spirometry reports did not include information on race/ethnicity that would have allowed classification of spirometric abnormality. Despite these limitations, this surveillance effort identified persons with abnormal declines that may have been occupational and merit individual follow-up.

Some lessons from these surveillance data for anticipated regulation follow. Medical surveillance with serial spirometry can identify workers at likely risk for irreversible lung disease related to diacetyl or its alpha-diketone substitutes. Minimizing economic impact for employers and workers would require improving spirometric quality and interpretation of serial declines. Companies can motivate the provision of good quality spirometry by making contracts under which payment is limited to tests which meet ATS quality goals (A or B grades) and under which lesser quality tests must be repeated within 1 month. Precedent exists for withholding payment to health care providers to motivate improved performance [Milstein, 2009]. Medical providers (or companies) can stipulate that technologists are certified for performance (in contrast to simply attending a certified course) and provide quality assurance for the spirometry services under their direction. These changes may be more likely to occur with independent audits of spirometry quality. Spirometry frequency might be tailored to risk factors for fixed obstructive lung disease and abnormal declines and by company surveillance experience. Provision for transfer of spirometry records when company medical providers change should be made in compliance with medical record privacy regulations. Medical surveillance in this industry is warranted because abnormal declines in FEV_1 in workers with occupational risk factors occurred in the period during which the flavoring companies were participating in an industry-wide prevention program with focused attention to respiratory protection and progressive increase in use of engineering controls. The substitution of other alpha-diketones, such as 2,3-pentanedione, for diacetyl is unlikely to lower risk [Flake et al., 2010; Hubbs et al., 2010; Morgan et al., 2011], and material safety data sheets have not included information about the presence of substitutes or their possible hazards based on structure–activity relationships [Day et al., 2011]. Worker participation in medical surveillance may depend on minimizing economic impact of medical evaluations, salary support, job protection, and counseling regarding personal protective equipment, work practices, and circumstances of restriction from further exposure. The changing nature of flavoring ingredients and improvement in engineering controls, personal protective equipment, and work practices may eventually make medical surveillance unnecessary. In the meantime, it will be a challenge to design appropriate regulations with procedures to evaluate their continuing need.

ACKNOWLEDGMENTS

The authors thank Len Welsh, former Cal/OSHA Chief, and his staff, especially Kelly Howard and Dan Leiner, for their collaboration in this work, and the participating workers, companies, and medical providers. Egils Kronlins assisted with data management. No authors had conflicts of interest (such as financial interests, relationships, or affiliations) relevant to the subject of this study. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

REFERENCES

AARC (American Association for Respiratory Care). 2010. AARC online spirometry certificate. http://www.aarc.org/osc/ (accessed 4/9/10).

Borg BM, Hartley MF, Fisher MT, Thompson BR. 2010. Spirometry training does not guarantee valid results. Respir Care 55(6):689–694.

California Department of Public Health, Occupational Health Branch. 2007. Medical surveillance for flavorings-related lung disease among flavor manufacturing workers in California [cited 2009 January 13]. Available from: http://www.cdph.ca.gov/programs/ohb/Documents/flavor-guidelines.pdf.

Day G, Le Bouf R, Grote A, Pendergrass S, Cummings K, Kreiss K, Kullman G. 2011. Identification and measurement of diacetyl substitutes in dry bakery mix production. J Occup Environ Hyg 8:93–103.

Dykstra BJ, Scanlon PD, Kester MM, Beck KC, Enright PL. 1999. Lung volumes in 4,774 patients with obstructive lung disease. Chest 115(1):68–74.

Enright PL, Beck KC, Sherrill DL. 2004. Repeatability of spirometry in 18,000 adult patients. Am J Respir Crit Care Med 169:235–238.

Enright PL, Skloot GS, Cox-Ganser JM, Udasin IG, Herbert R. 2010. Quality of spirometry performed by 13,599 participants in the World Trade Center worker and volunteer medical screening program. Respir Care 55(3):303–309.

Flake GP, Kirby PJ, Price HC, Gage PS, Kelly FL, Palmer SM, Foley JF, Morgan DL. 2010. Bronchiolitis obliterans-like lesions in rats treated with diacetyl, acetoin, or acetyl propionyl by intratracheal instillation. Toxicol Sci 114(S-1):316.

Hankinson JL, Odencrantz JR, Fedan KB. 1999. Spirometric reference values from a sample of the general U.S. population. Am J Respir Crit Care Med 159(1):179–187.

Harber P, Levine J, Bansal S. 2009. How frequently should workplace spirometry screening be performed? Optimization via analytic models. Chest 136(4):1086–1094.

Hnizdo E, Sircar K, Yan T, Harber P, Fleming J, Glindmeyer HW. 2007. Limits of longitudinal decline for the interpretation of annual changes in FEV_1 in individuals. Occup Environ Med 64:701–707.

Hnizdo E, Glindmeyer HW, Petsonk EL. 2010. Workplace spirometry monitoring for respiratory disease prevention: A methods review. Int J Tuberc Lung Dis 14(7):796–805.

Hubbs AF, Moseley AE, Goldsmith WT, Jackson MC, Kashon ML, Battelli LA, Schwegler-Berry D, Goravanahally MP, Frazer D, Fedan JS, Kreiss K, Castranova V. 2010. Airway epithelial toxicity of the flavoring agent, 2,3-pentanedione. Toxicol Sci 114(S-1):319.

Kim TJ, Materna BL, Prudhomme JC, Fedan K, Enright PL, Sahakian N, Windham G, Kreiss K. 2010. Industry-wide medical surveillance of California flavor manufacturing workers: Cross-sectional results. Am J Ind Med 53:857–865.

Kreiss K. 2007. Flavoring related bronchiolitis obliterans. Curr Opin Allergy Clin Immunol 7:162–167.

Leuppi JD, Miedinger D, Chhajed PN, Buess C, Schafroth S, Bucher HC, Tamm M. 2010. Quality of spirometry in primary care for case finding of airway obstruction in smokers. Respiration 79:469–474.

Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CPM, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J, for the ATS/ERS Task Force. 2005. Standardisation of spirometry. Eur Respir J 26(2):319–338.

Milstein A. 2009. Ending extra payment for ''never events''—Stronger incentives for patients' safety. N Engl J Med 360(23a):2388–2390.

Morgan DL, Jokinen MP, Price HC, Gwinn WM, Palmer SM, Flake GP. 2011. Bronchial and bronchiolar fibrosis in rats exposed to 2,3-pentanedione vapors: Implications for bronchiolitis obliterans in humans. Toxicol Path 2011:1–18. DOI: 10.1177/0192623311431946.

National Institute for Occupational Safety Health. 2010. Spirometry in the occupational setting: Spirometry longitudinal data analysis (SPIROLA) software. Atlanta, GA, USA: Centers for Disease Control and Prevention, NIOSH, Version 3.0, updated. November 2010. Available from: http://www.cdc.gov/niosh/topics/spirometry/spirola.html. Accessed May 2011.

SAS. 2004. SAS/STAT user's guide, version 9.1 for windows. Cary, NC: SAS Institute, Inc.

Townsend MC, Occupational and Environmental Lung Disorders Committee. 2011. Spirometry in the occupational health setting—2011 update. J Occup Environ Med 53:569–584.

U.S. Code of Federal Regulations. 2002. Standards for privacy of individually identifiable health information. Title 45, Part 164.512(b).

Wang ML, Petsonk EL. 2004. Repeated measures of FEV_1 over six to twelve months: What change is abnormal? J Occup Environ Med 46(6):591–595.

Wang ML, Avashia BH, Petsonk EL. 2006. Interpreting periodic lung function tests in individuals: The relationship between 1- to 5-year and long-term FEV_1 changes. Chest 130(2):493–499.