

# Correlates of Fitness for Duty in Hazardous Materials Firefighters

Stefanos N. Kales, MD, MPH,<sup>1,2,3,\*</sup> Jon M. Aldrich, BA,<sup>1</sup> Gerry N. Polyhronopoulos, BS,<sup>1</sup> Edward O. Leitaio, MD,<sup>1,2</sup> David Artzerounian, MD,<sup>4,5</sup> Thomas H. Gassert, MD, MSc,<sup>3,6,10</sup> Howard Hu, MD, ScD, MPH,<sup>3,6,7</sup> Karl T. Kelsey, MD, MOH,<sup>6,7,8</sup> Charles Sweet, MD, MPH,<sup>5,9</sup> and David C. Christiani, MD, MPH, MS,<sup>3,6,10</sup>

**Background and Methods** From a statewide medical examination program, we identified firefighters who were deemed unfit for duty by attending physicians (ATTENDING FAIL,  $n = 9$ ) and those who would have been disqualified by the application of selected numerical criteria from the 1997 National Fire Protection Association (NFPA) guidelines (NFPA FAIL,  $n = 27$ ) and criteria from a Medical Workshop (WORK FAIL,  $n = 16$ ). The subjects who were unfit for duty or failed numerical criteria were compared with those who were fit for duty and passed all objective criteria (FIT group,  $n = 302$ ). All subjects were given an overall morbidity rating by a board certified internist. Comparisons on two surrogate measures of fitness,  $VO_2$  max predicted and predicted coronary heart disease (CHD) risk, were also performed.

**Results** We found a significant tendency towards worse results (e.g. higher blood pressure or lower spirometric function) among the three FAIL groups compared with the FIT group. The FAIL groups shared only a small overlap, however, with the firefighters with the highest morbidity ratings, lowest predicted  $VO_2$  max, and highest CHD risks. Increasing morbidity was associated with higher age, lower spirometric function, lower predicted  $VO_2$  max, increasing cholesterol, greater BMI, and higher predicted 10 year CHD risk.

**Conclusion** Although the presence of a single serious or poorly controlled condition may render an individual unfit for safe performance as a firefighter, examination of our cohort suggests that multiple risk factor models or overall clinical assessments are superior means of identifying firefighters with poor health status and increased CHD risk. Am. J. Ind. Med. 36:618–629, 1999. © 1999 Wiley-Liss, Inc.

**KEY WORDS:** firefighters; fitness for duty; morbidity; body mass index; pulmonary function; cardiac/coronary risk predictors

<sup>1</sup>The Cambridge Hospital, Cambridge

<sup>2</sup>Harvard Medical School, Boston

<sup>3</sup>Harvard School of Public Health, Department of Environmental Health (Occupational Health Program), Boston

<sup>4</sup>The Work Connection at Holyoke Hospital, Holyoke, MA

<sup>5</sup>Department of Family and Community Medicine, UMASS Medical Center, Worcester, MA

<sup>6</sup>Center for Occupational and Environmental Medicine, Massachusetts Respiratory Hospital, Braintree, MA

<sup>7</sup>Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston

<sup>8</sup>Harvard School of Public Health, Department of Environmental Health (Occupational Health Program), Department of Cancer Cell Biology, Boston

<sup>9</sup>UMASS Health System, Marlborough Hospital, Marlborough, MA

<sup>10</sup>Pulmonary/Critical Care Unit, Massachusetts General Hospital/Harvard Medical School, Boston

Contract grant sponsor: NIOSH; Contract grant number: 1 K01 OH00156; Contract grant sponsor: NIH; Contract grant numbers: 1K07ES00266, ES00002 and ES05947; Contract grant sponsor: NORA; Contract grant number: 1 R01 OH03729.

\*Correspondence to: Stefanos N. Kales, MD, MPH, Occupational Medicine, Department of Medicine, The Cambridge Hospital, 1493 Cambridge Street, Cambridge, MA 02139. E-mail: stefokali@aol.com

Accepted 26 July 1999

## INTRODUCTION

There are an estimated one million persons in the United States involved in firefighting [Morse et al., 1992]. The 1997 National Fire Protection Association (NFPA) Standard Medical Requirements for firefighters [NFPA 1582, 1997] recommend medical examinations for all of these individuals. Additionally, an increasing number of communities have developed hazardous materials (hazmat) teams composed of specially-trained firefighters to respond to chemical spills, fires, and accidents. The OSHA standard on Hazardous Waste Workers [29 CFR.1910.120, 1989] requires medical surveillance for these hazmat firefighters and other potentially exposed workers. In Massachusetts, new firefighter candidates must now pass a physical abilities test, preplacement medical examination, and periodic medical examinations every two years thereafter. Most current incumbent firefighters in Massachusetts, however, do not undergo any required periodic medical testing.

Despite widespread recommendation and practice, it is unknown if medical examination programs can decrease morbidity and mortality, or if they correctly select persons who are medically capable or incapable of firefighting. Most surveillance tests and fitness for duty criteria are currently advocated on theoretical grounds, and limited scientific evidence is available to support their use in medical surveillance examinations and fitness for duty determinations [Kales et al., 1997, 1998]. Fitness guidelines (e.g., NFPA 1582) traditionally review at what point, or in the presence of what manifestations, specific conditions or diseases become "incompatible with the safe performance of essential job functions..." [Samo and Bogucki, 1999]. Even less data can be used to support or refute various cutoff values beyond which point a firefighter is suggested to become unfit for duty [Kales et al., 1998; Kales and Christiani, 1999]. No "gold standard" for determining fitness or for validating proposed fitness criteria exists.

Given the absence of a "gold standard" or a prospective study of outcomes, surrogate measures of fitness could be considered. The NFPA [NFPA 1582, 1997] regards the following to be among the essential capabilities of a firefighter: the ability to "... make rapid transitions from rest to near maximal exertion without warm-up periods"; to wear about 50 pounds of personal protective equipment, and at the same time to "... perform physically demanding work . . . at near maximal heart rates for prolonged periods of time...". Therefore, an adequate aerobic capacity or exercise tolerance might be one surrogate of fitness for duty.

Another goal of fitness determination is the identification of individuals at high risk for incapacitation and death during firefighting. Fatality data collected by the NFPA show that the on-duty death rate is over six times higher for firefighters over 60 years of age and three times higher for 50–59 year olds than for those 20–39 years old [Washburn,

1998]. The NFPA has also shown that heart attacks are consistently the most frequent cause of death among firefighters, accounting for 45% of US fatalities during the period 1977–1996 [NFPA, 1998]. Over 90% of these deaths occurred in individuals with known heart disease or significant coronary heart disease risk factors. Fifty per cent had prior myocardial infarctions or bypass surgery, another 32% had severe coronary atherosclerosis, and an additional 12% had diabetes or hypertension [Washburn, 1998]. The absolute number of heart disease deaths, however, is about 40–45 per year [NFPA, 1998; Washburn, 1998], so the overall risk of death is low. Therefore, it is unclear how to best exclude those at the highest risk while avoiding age discrimination and without disqualifying a large number of firefighters from duty who, despite risk factors, will perform without significant problems.

We previously examined the clinical yields of various examinations, and the practical implications of applying selected NFPA and other fitness criteria to a cohort of hazardous materials firefighters [Kales et al., 1998]. We also compared these suggested fitness criteria to the attending physicians' determinations of fitness for duty. Because no "gold standard" exists for determining fitness or for validating proposed fitness criteria, we could not test the sensitivity, specificity, or predictive values of the fitness criteria or the physicians' clinical fitness determinations.

Although the hypothesis might be considered intuitive, we are unaware of any investigations that demonstrate that firefighters who are found "unfit for duty" or who fail to meet selected NFPA or other proposed criteria are actually less healthy than their "fit" colleagues as measured by objective testing. Likewise, it has not been shown that such persons are incapable of safely performing as firefighters. Therefore, in this study, we compared the health profiles of firefighters who failed fitness for duty determinations by attending physicians and those who failed various objective criteria with the profiles of firefighters who were judged fit for duty and passed all objective criteria. We included comparisons on two surrogate measures of fitness,  $VO_2$  max predicted and predicted coronary heart disease (CHD) risk, as both a poor aerobic capacity and/or high risk of cardiac incapacitation may impair the safe and successful performance of essential firefighting job functions. We also examined the same health/fitness profiles in association with low, moderate, and high clinical morbidity ratings.

## METHODS

### Subjects

The subjects were all members ( $n = 340$ ) of six regional hazardous materials response teams of the Commonwealth of Massachusetts as well as regular municipal firefighters for local departments. The population included four women

(1%). The team members' ages ranged from 21 to 58 with a mean of 39 and a standard deviation of 6.9. Some had undergone previous surveillance examinations, while for others, these were their baseline medical examinations after joining the teams.

### Baseline and Periodic Medical Examinations

These examinations were performed for the dual purposes of medical surveillance of hazardous materials response duty and determination of fitness for the state hazardous materials teams. These fitness determinations for state duty do not apply to firefighters' work status/capacity with their municipal (non-hazmat) fire departments.

Medical surveillance examinations for all subjects were performed at three hospitals: Massachusetts Respiratory, Marlborough or Holyoke Hospital in 1996 or 1997 in the first year of a statewide surveillance program. All examinations were conducted in a similar fashion: similar history forms, laboratory tests and standardized physical examination procedures were utilized. The components of the examinations were in accordance with both the OSHA standard on Hazardous Waste Workers (29 CFR.1910.120, 1989) and the NFPA 1582 medical requirements for firefighters [NFPA 1582, 1997]. Determinations of fitness were left up to the judgment of the attending physicians. No predetermined fitness criteria were applied. The attending physicians were all board certified in occupational medicine, and all except one were also board certified in internal medicine.

Examinations included a detailed medical, smoking, and environmental/occupational history tailored to emergency responders; physical examination; visual and audiometric testing; routine laboratory tests (complete blood count, blood urea nitrogen, creatinine, alkaline phosphatase, AST and ALT, and urinalysis); and spirometry. All individuals also underwent a minifitness evaluation: this test required a firefighter to ascend and descend two flights of stairs while wearing a respirator with pre- and post-test assessment of vital signs. Electrocardiogram, posterior-anterior chest roentgenogram, and RBC cholinesterase activity were also done as baselines if the 1996/1997 exam was the individual's baseline examination or the test had not been done previously.

Summary results of each firefighter's examination, including the attending physician's determination of fitness for duty, were transferred to a statewide computerized medical records repository at the Massachusetts Respiratory Hospital. The body mass index (BMI) in kg/m<sup>2</sup> of each subject was calculated from his/her height in inches and weight in pounds using the following formula [Pi-Sunyer, 1992]:

$$\text{BMI} = 703.1 \times \text{weight in pounds}/(\text{height in inches})^2$$

### Fitness Criteria

Various methods and criteria for determining "fitness" were selected for study. These included: (1) the fitness for duty determinations by the examining attending occupational medicine physician; and (2) simulated *post hoc* fitness determinations based on (A) selected guidelines from NFPA 1582 [1997]; and (B) criteria agreed upon by the investigators and examining physicians at a medical "workshop" meeting.<sup>1</sup> Those determined to be unfit by the attending physician are identified as the "ATTENDING FAIL" group. The subjects who failed either the proposed NFPA or the workshop criteria during the *post hoc* review were identified as the "NFPA FAIL" and "WORKSHOP FAIL" groups, respectively. Those firefighters who satisfied all of the selected NFPA and medical workshop criteria and were determined to be fit by an attending physician were classified as the "FIT" group. Table I shows various cutoffs for resting blood pressure, post minifitness test blood pressure, and visual and audiometric deficits based on NFPA and workshop consensus criteria for determining that a firefighter is unfit for duty.

### Determination of the Severity of Visual and Acoustic Deficits

All firefighters with corrected far visual acuity worse than 20/30 in one or both eyes were selected for further review. Binocular acuities were obtained from their examination charts for baseline or hazmat clearance at the 3 participating hospitals and compared to the study criteria.

The audiograms for all firefighters (n = 124) judged abnormal by the attending physician were obtained and systematically reviewed against the NFPA and medical workshop criteria. Those who were judged normal or "borderline" were assumed to pass the study criteria.

### Overall Morbidity Rating ("fitrank")

The face problem sheet of each subject was reviewed by a board-certified internist (EOL). EOL had never examined any of the firefighters and was blinded to the subjects' fitness outcomes. The face problem sheet lists gender, age, height and weight, medical problems, medications, allergies, immunizations, and smoking history. All information listed on the face sheets except fitness outcomes was available to the reviewer. EOL rated each subject on a scale of 1–10 for increasing morbidity. A rating of 1 was considered most

<sup>1</sup> The workshop took place in January 1997. Actual examination results were discussed, and the proposed NFPA criteria were reviewed with state officials, firefighter representatives, and the examining physicians. Participants in developing the workshop criteria were Drs. Kales and Christiani, SNK, DCC and the examining physicians Drs. Artreronian, Gassert, Hu, Kelsey, and Sweet. DA, TG, HH, KTK, and CS.

**TABLE I.** Objective Criteria for Exclusion from Fitness for Duty

Source	Resting blood pressure	Minifitness evaluation	Visual acuity	Audiometry	Pulmonary function
1997 revision of NFPA 1582	SBP > 179, DBP > 99	NA	Corrected binocular vision worse than 20/30	Hearing deficit in <i>unaided</i> worst ear: >25 dB in 3 of the 4 frequencies: 500, 1000, 2000, 3000 Hz or; >30 dB in either 500, 1000, or 2000 Hz <i>and</i> average >30 dB for: 500, 1000, 2000, 3000 Hz	No cutoff given
Medical workshop	DBP > 109	Post exertion SBP > 219 or DBP ≥ 109	Corrected binocular vision worse than 20/40	Same as NFPA, but for best <i>aided</i> ear.	No cutoff given

Please note, Table I presents objective study criteria and does not refer exclusively to our cohort of hazardous materials firefighters.

healthy and 10 least healthy. This variable was labeled FITRANK. For the purpose of comparisons, low, moderate and high morbidity groups were defined as follows: low, FITRANK of 1–3; moderate, FITRANK of 4–6, and high, FITRANK of 7–10.

### Calculation of Estimated Aerobic Capacity (VO<sub>2</sub> max)

We used the mean age, BMI, and VO<sub>2</sub> max for each 10 year age grouping from three studies [Lemon et al., 1977; Horowitz and Montgomery, 1993; Kilbom, 1980] and the overall means for VO<sub>2</sub> max, mean age, and mean BMI from two other studies [Davis et al., 1982; Boreham et al., 1994], to derive a regression equation for VO<sub>2</sub> max based on age and BMI. The equation is shown here:

$$\text{VO}_2 \text{ max} = 94.02 - 0.44(\text{age}) - 1.52(\text{BMI});$$

Multiple R = 0.961; Squared Multiple R = 0.923; Age ( $P < .0001$ ); BMI ( $P = .006$ )

### Predicted Risk of Coronary Heart Disease

Each firefighter's risk of coronary heart disease over a period of 10 years was calculated from data from the Framingham study [Wilson et al., 1998]. The Framingham study formulated sex-specific coronary prediction algorithms according to age, diabetes, smoking, Joint National Committee blood pressure categories, National Cholesterol Education Program total cholesterol and LDL cholesterol categories, and HDL cholesterol. The algorithm predicts an average risk based on the CHD experience of a sample of 2489 men and 2856 women 30–74 years old at the point of their baseline examination between 1971 and 1974. We selected this prediction model because our cohort of predominantly white, suburban Massachusetts firefighters

is similar to the white subjects drawn from a community in the western suburbs of Boston.

The CHD score sheet assigns a point value for each of the six categories and the predicted average risk is determined from the point total. Because of the small number of women ( $n = 4$ ) in our cohort and the fact that the prediction equations for men and women are different, we excluded females from the CHD risk analysis. We assigned points in each of the six categories for the 154 firefighters for whom all necessary data are available. It is important to note the total cholesterol categories in the Framingham study were based on a blood test after an overnight fast. The total cholesterol values for the firefighters in our cohort were not fasting values, and this may slightly skew the results toward greater risk. On the other hand, we did not have direct information for HDL-cholesterol. Because the average cholesterol was over 225 mg/dL, we assumed that the average HDL value for our cohort would be lower than 44 mg/dL. Therefore, based on a point system in which <35 mg/dL = 2 points, 35–44 mg/dl = 1 point, 45–49 and 50–59 both = 0 points, and  $\geq 60 = -2$  points, we conservatively assigned a value of 1 point to each firefighter for the HDL category.

To validate our modified prediction model, we compared the mean 10 year CHD risks that we calculated with the average risks of the male subjects in the Framingham Study for five year age groupings. This is shown in Table II. The mean 10 year CHD risks for our male firefighter cohort correspond closely with those in the Wilson et al. [1998] Framingham study.

We derived each firefighter's estimated risk of coronary disease over a five year period from an earlier Framingham study [Anderson et al., 1991] of predicted CHD risk with a similar predictive model to the Wilson study. The Anderson et al. [1991] study's risk prediction equation provides estimates of both five and ten year risk. Given the similar

**TABLE II.** Predicted Coronary Heart Disease Risk of Firefighter Age Groups (Males Only), Hazardous Materials Firefighters, Massachusetts

Age group	N	Mean predicted 10 year CHD risk	Average 10 year CHD risk (Framingham study)
<30	6	3.0 ± 0.9%	NA
30–34	35	4.3 ± 1.5%	3%
35–39	36	5.1 ± 2.4%	5%
40–44	36	7.1 ± 2.9%	7%
45–49	30	10.8 ± 3.5%	11%
50–54	10	14.4 ± 5.4%	14%
55–59	1	13.0	16%
Total	154	7.1 ± 4.2%	NA

ANOVA: Between age groups (Combined).

Sum of squares = 1487.3; df = 6; mean square = 247.9; F = 31.1; P &lt; 0.001.

predictive capabilities of the two algorithms, the ten year risk values derived from the Wilson et al. [1998] study approximate closely ten year risk values reported in the Anderson et al. [1991] study. Each firefighter's approximate five year risk was determined by selecting the five-year probabilities that correspond to each ten year probability in the Anderson et al. [1991] study.

### Statistical Analyses

We hypothesized *a priori* that the various unfit groups would be older, have higher mean resting blood pressures, serum cholesterol, BMI, and overall morbidity ratings, and lower pulmonary function than the FIT group. Likewise, we hypothesized *a priori* that the groups with higher overall morbidity ratings would be older, have higher resting blood pressures, cholesterol levels, BMIs, and lower pulmonary function than the groups with lower overall morbidity ratings.

We used independent two-tailed t-tests to determine if the variable means of a FAIL group differed from the FIT group. Determinations of statistical significance were based on the assumption of separate variances between groups. Proportions were compared using the standard (Pearson's) chi-square test with the exception of comparisons where one or more cells had a frequency of less than five. Fisher's exact test was used to compare these proportions.

To compare groups with respect to differences on multiple variable means simultaneously, we used the sign test [Hines and Montgomery, 1980]. Dependent, closely related variables (e.g. systolic and diastolic blood pressure, and FVC % and FEV1%) were treated as one variable and variables calculated from independent variables (VO<sub>2</sub> max predicted, and predicted 10 year CHD risk) were not considered. This avoided counting differences on non-independent variables twice.

## RESULTS

### Fitness for Duty Determinations

A total of 331 (97%) firefighters were determined to be fit for duty by the attending physicians, and nine (3%), all males, were judged unfit. Seven failed, at least in part, due to their elevated resting and/or post exercise blood pressures. Two of these seven also were cited for abnormal liver function tests. One was unfit due to an orthopedic injury. The rationale for the ninth individual's failure was not explicitly present in the central database.

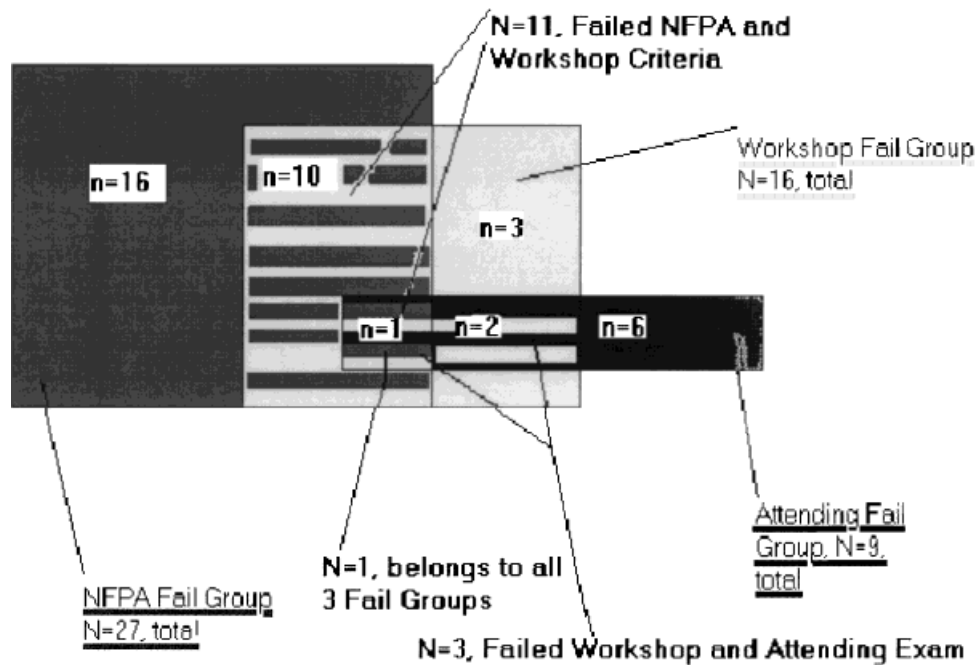
A total of 27 firefighters, the NFPA FAIL group, failed one or more of the selected NFPA criteria. A total of 16 firefighters, the WORKSHOP FAIL group, failed one or more of the medical workshop criteria. Table III summarizes the performance failures of the entire cohort on selected NFPA and medical workshop criteria. A total of 302 firefighters, the FIT group, satisfied all of the selected NFPA and medical workshop criteria and were determined to be fit by an attending physician. Figure 1 illustrates the overlap of the various FAIL groups.

### Physiological/Metabolic/Morbidity Profiles of Fit vs. Various Fail Groups

Table IV describes the metabolic, physiological, and morbidity profiles of the FIT, NFPA FAIL, WORKSHOP FAIL, and ATTENDING FAIL groups. The FIT group was younger than all three FAIL groups and significantly younger than the NFPA and WORKSHOP FAIL groups (both  $P < .05$ ). Resting blood pressures were also higher for all three FAIL groups when compared to the FIT group. Both the systolic and diastolic blood pressures of the ATTENDING FAIL group were significantly higher (both  $P < .01$ ). The systolic blood pressure

**TABLE III.** Performance Failures of Total Cohort on Study Criteria, Hazardous Materials Firefighters, Massachusetts

Source	Resting blood pressure	Minifitness evaluation	Visual acuity	Audiometry	Total
1997 revision of NFPA 1582	7/340 (2.0%)	NA	5/335(1%)	16/330(5%)	27
Medical Workshop	1/340 (0.3%)	7/336 (2%)	1/335 (0.3%)	8/330 (2%)	16

**FIGURE 1.** Overlap of various FAIL groups, Hazardous Materials Firefighters, Massachusetts.

of the WORKSHOP FAIL group was also significantly higher ( $P < .05$ ).

The mean BUN of the FAIL groups were higher than that of the FIT group and was significantly higher for the ATTENDING FAIL group ( $P < .05$ ). Other significant differences were observed between the WORKSHOP FAIL and FIT groups for FIT RANK (overall morbidity rating) ( $P < .05$ ), and the mean predicted  $VO_2$  max ( $P < .05$ ).

We found a significant tendency towards slightly worse results (e.g. higher blood pressure or lower spirometric function) for the profile variables for the NFPA FAIL, WORKSHOP FAIL, and ATTENDING FAIL groups compared with the FIT group. In a Sign Test of the 35 possible variable comparisons<sup>2</sup>, there were 4 ties, while 28 of 31 variable means were worse for the FAIL groups when compared to the FIT group ( $P < .01$ ). Moreover, another Sign Test considered only those variables<sup>3</sup> for which we had an a priori hypothesis that the FAIL groups would perform worse

than the FIT group. In this comparison of 20 variable means, there was one tie and 19 of the 19 total comparisons ( $P < .01$ ) between variable means were worse for the FAIL groups compared to the FIT group.

### Physiological/Metabolic/Morbidity Profiles of Firefighters Classified by Overall Morbidity Rankings ("fitrank")

Table V describes the metabolic, physiological, and morbidity profiles of the cohort grouped by low, moderate, and high overall morbidity rankings. The low morbidity group was significantly younger than both the moderate and high morbidity groups (both  $P < .01$ ). Increasing morbidity was associated with lower spirometric function: both the mean percent predicted FVC and FEV1 were significantly lower for the high morbidity group than the low morbidity group ( $P < .05$  and  $P < .01$ , respectively).

The increasing morbidity was associated with the increasing cholesterol and BMI. The mean cholesterol was significantly higher for the moderate and high morbidity groups ( $P < .05$ ). The mean BMI of the low group was

<sup>2</sup> Systolic and diastolic blood pressure are considered as one variable, as are FVC % and FEV1%.

<sup>3</sup> Age, SBP, DBP, FVC %, FEV1%, cholesterol, BMI, and Overall Morbidity Rating.

**TABLE IV.** Physiological, Metabolic, and Morbidity Profiles of Fit vs. Various Fail Groups, Hazardous Materials Firefighters, Massachusetts

Variable	Fit by all N = 302	NFPA fail N = 27	Workshop fail N = 16	Unfit per MD N = 9
Age	39.0 ± 6.8 N = 287	43.3 ± 7.5 N = 25; P = .011	44.3 ± 7.1 N = 15; P = .014	40.7 ± 8.4 N = 9; P = .577
SBP	122 ± 12 N = 302	127 ± 21 N = 27; P = .181	135 ± 21 N = 16; P = .022	134 ± 11 N = 9; P = .008
DBP	78 ± 8 N = 302	82 ± 15 N = 27; P = .202	85 ± 14 N = 16; P = .062	90 ± 9 N = 9; P = .005
FVC%	103 ± 13 N = 298	103 ± 13 N = 25; P = .816	100 ± 13 N = 14; P = .398	96 ± 15 N = 9; P = .189
FEV1%	103 ± 13 N = 298	101 ± 14 N = 25; P = .381	99 ± 12 N = 14; P = .192	96 ± 13 N = 9; P = .157
ALKPHOS	82 ± 21 N = 298	85 ± 24 N = 27; P = .499	84 ± 21 N = 16; P = .691	91 ± 31 N = 9; P = .427
CHOL	226 ± 40 N = 173	237 ± 39 N = 19; P = .234	246 ± 43 N = 8; P = .226	Insufficient data
BUN	16 ± 4 N = 301	17 ± 4 N = 27; P = .094	18 ± 5 N = 16; P = .143	19 ± 3 N = 9; P = .026
CR	1.1 ± 0.2 N = 301	1.1 ± 0.1 N = 27; P = .378	1.1 ± 0.2 N = 16; P = .593	1.1 ± 0.2 N = 9; P = .517
AST	25 ± 10 N = 299	24 ± 8 N = 27; P = .823	24 ± 5 N = 16; P = .882	26 ± 0 N = 9; P = .672
ALT	36 ± 20 N = 299	39 ± 18 N = 27; P = .362	35 ± 12 N = 16; P = .788	45 ± 25 N = 9; P = .285
BMI	28.9 ± 4.1 N = 296	28.9 ± 3.3 N = 26; P = .980	30.4 ± 3.2 N = 15; P = .088	30.8 ± 3.1 N = 8; P = .119
FITRANK	3.7 ± 1.3 N = 302	3.9 ± 1.4 N = 27; P = .550	4.4 ± 1.2 N = 16; P = .039	3.8 ± 0.7 N = 9; P = .683
VO <sub>2</sub> max predicted	33 ± 7 N = 281	31 ± 7 N = 24; P = .217	28 ± 6 N = 14; P = .017	30 ± 5 N = 8; P = .118
Predicted 10 year CHD risk	6.8 ± 4.1% N = 138	9.3 ± 4.6% N = 15; P = .063	8.0 ± 3.1% N = 6; P = .393	Insufficient data

significantly lower than the means for both the moderate ( $P < .01$ ) and the high morbidity groups ( $P < .05$ ). The mean predicted VO<sub>2</sub> max of the moderate and high morbidity groups was significantly lower than that of the low morbidity group (both  $P < .01$ ). Finally, the predicted 10 year CHD risks were significantly higher for the moderate and high morbidity groups ( $P < .001$  and  $P < .005$ , respectively) compared to the low morbidity group.

When all the profile means of the low morbidity group were compared with those of the moderate and high morbidity groups, the Sign Test did not reach statistical significance for either group. For the moderate morbidity group, 7 of 8 means were worse, with 3 ties ( $P = .10$ ). For the high morbidity group, 7 of 9 were worse, with 2 ties ( $P \geq .10$ ). However, when only those variables<sup>4</sup> for which

we had an *a priori* hypothesis were considered, the Sign Test was significant in both cases. Both the moderate and high morbidity groups performed worse on 6 of 6 means compared to the low morbidity group (both  $P < .05$ ).

The prevalence of ever smoking and current smoking increased with increasing morbidity. This was expected, as the rating internist was not blinded to smoking status and used this information in rating subjects. In the low morbidity group, 14% (20/138) were current or former smokers compared to 50% (65/131) of the moderate ( $P < .01$ ) and 78% (7/9) of the high morbidity group ( $P < .00$ , Fisher's Exact Test). No current smokers were found in the low morbidity group (0/153) compared to 20% (29/142) of the moderate ( $P < .01$ ) and 44% (4/9) of the high morbidity group ( $P < .01$ , Fisher's Exact Test).<sup>5</sup>

<sup>4</sup> Age, SBP, DBP, FVC %, FEV1%, cholesterol, BMI, and Overall Morbidity Rating.

<sup>5</sup> The N's for smoking prevalence vary due to non-uniform answers; i.e. never smoker vs. non-smoker.

**TABLE V.** Physiological, Metabolic, and Morbidity Profiles of Firefighters Classified by Overall Morbidity Ratings (“FITRANK”), Hazardous Materials Firefighters, Massachusetts

Variable	Fitrank 1–3 N = 173	Fitrank 4–6 N = 158	Fitrank 7–10 N = 9
Age	37.4 ± 6.7 N = 162	40.9 ± 6.5 N = 152; <i>P</i> = .000	49.4 ± 4.4 N = 9; <i>P</i> = .000
SBP	122 ± 13 N = 173	123 ± 13 N = 158; <i>P</i> = .673	127 ± 17 N = 9; <i>P</i> = .438
DBP	79 ± 9 N = 173	79 ± 9 N = 158; <i>P</i> = .490	79 ± 10 N = 9; <i>P</i> = .972
FVC %	104 ± 14 N = 168	102 ± 12 N = 157; <i>P</i> = .130	92 ± 12 N = 9; <i>P</i> = .021
FEV1 %	104 ± 14 N = 168	102 ± 12 N = 157; <i>P</i> = .096	89 ± 10 N = 9; <i>P</i> = .002
ALKPHOS	82 ± 20 N = 169	83 ± 22 N = 158; <i>P</i> = .736	89 ± 30 N = 9; <i>P</i> = .539
CHOL	220 ± 43 N = 105	235 ± 34 N = 82; <i>P</i> = .011	247 ± 22 N = 7; <i>P</i> = .019
BUN	16 ± 4 N = 172	16 ± 4 N = 158; <i>P</i> = .165	16 ± 4 N = 9; <i>P</i> = .863
CR	1.1 ± 0.2 N = 172	1.1 ± 0.2 N = 158; <i>P</i> = .014	1.1 ± 0.1 N = 9; <i>P</i> = .269
AST	25 ± 12 N = 171	25 ± 8 N = 157; <i>P</i> = .724	24 ± 7 N = 9; <i>P</i> = .818
ALT	37 ± 23 N = 171	36 ± 18 N = 157; <i>P</i> = .427	30 ± 14 N = 9; <i>P</i> = .194
BMI	27.9 ± 3.5 N = 171	30.0 ± 4.3 N = 153; <i>P</i> = .000	31.2 ± 3.3 N = 9; <i>P</i> = .015
Predicted 10 year CHD risk	5.3 ± 2.6% N = 80	8.1 ± 3.9% N = 67; <i>P</i> = .000	16.3 ± 6.2% N = 7; <i>P</i> = .003
VO <sub>2</sub> max predicted	35 ± 6 N = 160	31 ± 7 N = 147; <i>P</i> = .000	25 ± 6 N = 9; <i>P</i> = .001

The likelihood of failing the attending physician’s fitness determination and/or any one of the study criteria was essentially the same across morbidity rankings of 2–6, varying only from 10–13%. It differed only at the extremes: 0% (0/11) for those with a morbidity rating of 1, and 22% (2/9) for those with ratings of 7–10. Figure 2 shows the minimal overlap between the high morbidity group and the FAIL groups.

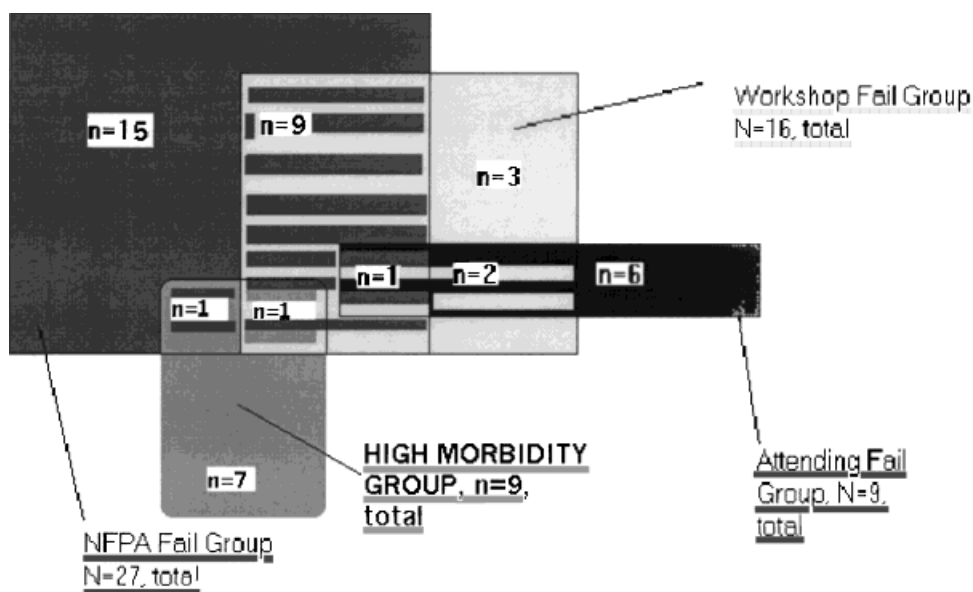
### Predicted Coronary Heart Disease Risk

An examination of the mean predicted CHD risk for 5 year age categories (Table II), discussed above in the Methods section, demonstrates a statistical association between increasing categorical age and mean predicted 10 year CHD risk (ANOVA between age groups, *P* < .001). Based on the NFPA recommendation that firefighters age 40 and older undergo exercise stress testing, we examined the

ability of this guideline to identify individuals with 5 year CHD risks  $\geq 5\%$  and  $\geq 3\%$ . These data are displayed in Figure 3.

Implementing a guideline of exercise stress testing for all firefighters aged 40 years or more would select for stress testing 40 of the 42 firefighters (95%) with a 5 year CHD risk of 5% or greater. This guideline, however, would also select 37 firefighters for stress testing who have a 5 year CHD risk that is less than 5%. For selecting firefighters with a 5 year CHD risk  $\geq 3\%$ , the age-based guideline is less sensitive: 59 of 74 (80%). For 5 year CHD risk  $\geq 3\%$ , the age-based guideline is more specific in that only 18 of the 77 (23%) subjects selected for testing by the age-based guideline would have had predicted 5 year CHD risk that is <3%.

The overlap of the low, moderate, and high morbidity groups with the subgroups with increased CHD risk is shown in Table VI. Although elevated CHD risks are more



**FIGURE 2.** Overlap between FAIL groups and physician-derived “high” morbidity group, Hazardous Materials Firefighters, Massachusetts.

**TABLE VI.** Overlap Among Morbidity Groups and 5 Year Coronary Heart Disease Risk Groups, Hazardous Materials Firefighters, Massachusetts

5 year coronary heart disease risk	<3% N (% of row total)	3 to <5% N (% of row total)	≥ 5% N (% of row total)
Low morbidity group N = 80	55 (69%)	15 (19%)	10 (12%)
Moderate morbidity group N = 67	25 (37%)	16 (24%)	26 (39%)
High morbidity group N = 7	0 (0%)	1 (14%)	6 (86%)

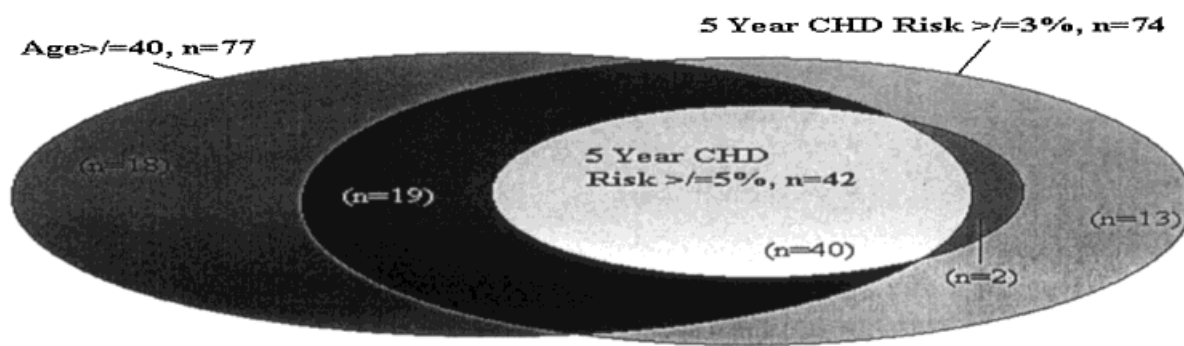
likely with increasing morbidity scores, substantial numbers of firefighters with low or moderate morbidity ratings also had 5 year CHD risks  $\geq 5\%$ .

**DISCUSSION**

In this study, we have demonstrated that firefighters who failed a physician’s fitness for duty determination, as well as those who failed selected numerical fitness criteria, on average had worse physiological and metabolic profiles than firefighters who passed the attending examination and all study criteria. On the other hand, these criteria did not tend to select firefighters with the highest morbidity scores, lowest predicted VO<sub>2</sub> max, and highest predicted CHD risks. The overlap of the FAIL groups with the high morbidity group is shown in Figure 2.

The converse criticism can be made for the NFPA age-based guidelines for stress testing (see Fig. 3). Since the NFPA also suggests stress testing for persons over 35 with

“one or more coronary artery disease risk factors [premature family history (less than age 55), hypertension, diabetes mellitus, cigarette smoking, and hypercholesterolemia (total cholesterol >240 or HDL cholesterol <35) [NFPA 1582, 1997]”, the sensitivities for identifying all persons with CHD risks at or above the cutoff values would be even higher than those shown in Figure 3. Despite the strong statistical association between 10 year CHD risk and increase in categorical age in this study and the documented increased on-duty death rates for older firefighters [Washburn, 1998], CHD risk is dependent also on other factors. Implementation of primarily age-based guidelines would lead to testing large numbers of firefighters with low CHD risks. Given that our data for the firefighter cohort for five year age intervals match closely with the Framingham data [Wilson et al., 1998] (see Table II), our investigation suggests that a multivariate risk factor selection protocol is the most accurate and cost efficient method of selecting firefighters for exercise stress testing.



Age	Five-Year CHD Risk		Totals:
	≥ 5%	< 5%	
≥ 40	40	37	77 (50%)
< 40	2	75	77 (50%)
<b>Totals:</b>	<b>42 (27%)</b>	<b>112 (73%)</b>	<b>154 (100%)</b>

Sensitivity = 95% (40/42)  
 Specificity = 67% (75/112)  
 Positive predictive value = 52% (40/77)  
 Negative predictive value = 97% (75/77)

Age	Five-Year CHD Risk		Totals:
	≥ 3%	< 3%	
≥ 40	59	18	77 (50%)
< 40	15	62	77 (50%)
<b>Totals:</b>	<b>74 (48%)</b>	<b>80 (52%)</b>	<b>154 (100%)</b>

Sensitivity = 80% (59/74)  
 Specificity = 78% (62/80)  
 Positive predictive value = 77% (59/77)  
 Negative predictive value = 80% (62/77)

**FIGURE 3.** Overlap between cohort subgroup ≥40 years of age and and coronary heart disease risk groups, Hazardous Materials Firefighters, Massachusetts.

In light of the NFPA data showing that heart attack is the predominant cause of death among firefighters, and the evidence that 90% of these fatalities occurred in firefighters with known CHD or CHD risk factors, our findings regarding CHD risk are particularly relevant. Moreover, the use of an evidence-based CHD prediction model is more likely to avoid violations of the Americans with Disabilities Act (ADA) because it would not identify firefighters for further testing on the basis of their age alone. It is still unclear, however, at what level of increased CHD risk the threshold for exercise stress testing should be set. Also unknown is how stress testing and subsequent evaluations of positive results would actually perform in distinguishing those at the highest risk of myocardial infarction and death. Such questions can only be answered by prospective studies.

Other findings that deserve further consideration regard the high morbidity group. Our data show that the predicted CHD risk is greater for both the moderate and high morbidity groups when compared to the low morbidity group. The high morbidity group had a 10 year CHD risk of 16.3%, while based on their age, an average risk of 12.6% is expected. This group was older, had a larger proportion of smokers, higher cholesterol, lower pulmonary function, more obesity (mean BMI >30), and a lower VO<sub>2</sub> MAX

predicted (mean VO<sub>2</sub> max 25 ± 6 mL/kg · min). A recent study [Hegmann and Tavis, 1998] directly measured aerobic capacity in firefighters. They found the same risk factors or closely related ones (increased age, smoking, lower % FVC, and adverse lipid profiles) to be predictors of an aerobic capacity < 30.0 mL/kg · min. In addition, other studies have consistently demonstrated that an increasing BMI is also inversely associated with aerobic capacity [Davis et al., 1982; Jette and Sidney, 1990].

The high morbidity group appears to be the least healthy 3% of our cohort. Our findings suggest that such individuals with multiple medical problems are likely to have both low exercise tolerances and elevated cardiac risks. Therefore, stress testing of firefighters high morbidity should be helpful in quantifying their work capacities and risks for cardiac incapacitation.

Despite known predictors of poor aerobic capacity, the appropriate minimum work capacity required of a firefighter is unclear. In a study of simulated firefighting, the mean VO<sub>2</sub> was 41.5 mL/kg · min during the most demanding operations. To support this work effort for a period of ten minutes requires a VO<sub>2</sub> max of 47.4 mL/kg · min [Gledhill and Jamnik, 1992]. Both Gledhill and Jamnik [1992] and Horowitz and Montgomery [1993] recommended that

firefighter applicants should have a minimum aerobic capacity of 45 mL/min · kg. The NFPA suggests a minimum exercise capacity of 37 mL/min · kg (10 METS) as one of the prerequisites for the clearance of persons with known coronary artery disease [NFPA 1582, 1997]. They do not require or suggest a minimum aerobic capacity criterion for firefighters in general.

In contrast to these recommendations, studies of actual firefighter cohorts have shown considerably lower mean aerobic capacities. Based on the present study, and six other studies of incumbent firefighters that found aerobic capacities to range between 33 and 42 mL/min · kg [Davis et al., 1982; Lemon and Hermiston, 1977; Horowitz et al., 1993; Kilbom, 1980; Boreham et al., 1994; Hegmann and Tavis, 1998], we can reasonably conclude that substantial proportions of these actual cohorts lack the proposed “minimum” exercise tolerances. One possible explanation is that those firefighters with greater aerobic capacities are generally assigned or self-select the most demanding firefighting operations. Alternatively, simulations and recommended criteria may overestimate the actual minimum requirements. Simulated tasks measure one’s ability to accomplish a task in one specific manner, while experienced firefighters may know less physically demanding ways to accomplish the same task. Such discrepancies emphasize the difficulty in reconciling physical testing standards with actual job performance. This is one of the reasons that NFPA 1582 was designed as a “medical standard” to identify medical conditions incompatible with the safe performance of essential job functions, but not a “physical standard” which would be designed to evaluate specific physical capabilities [Samo and Bogucki, 1999].

Several additional observations should be made regarding the selected numerical fitness criteria which were related to blood pressure, vision and hearing. Unlike blood pressure guidelines, the minimum hearing and vision criteria are advocated to reduce safety problems related to poor sensory input. One weakness of this study was the small number ( $n=10$ ) of firefighters who failed one or more of the proposed blood pressure guidelines and for whom we had incomplete data for cardiac risk factors. Therefore, we were unable to examine the CHD risks of those failing blood pressure guidelines independent of those who failed vision and hearing criteria.

Obviously, the appropriateness of minimum vision and hearing guidelines cannot be judged on the basis of selection for poor health, prediction of low aerobic capacity or increased CHD risk. The validation of vision and hearing criteria needs to come from prospective studies that document increased risks of adverse outcomes for firefighters who fail the said criteria. This cohort is being followed prospectively and should provide some information on whether firefighters who lack the minimum proposed criteria actually have an increased risk of adverse outcomes.

The fact that firefighters who failed the selected numerical fitness criteria have on average worse physiological and metabolic profiles probably relates to the association of increasing age with increasing blood pressure and declining vision and hearing. We did not control for age in this study because under the ADA, fitness determinations should not be made on the basis of age. In addition, being an outlier in one area might be associated with other health problems, and such individuals may be at an increased risk for several reasons. Again, the prospective studies should help clarify such risks.

This study has several limitations. First, its cross-sectional nature and lack of exercise stress testing do not allow us to directly validate the aerobic capacities and CHD risks that we predicted. Although we did not directly measure aerobic capacity, our prediction model is likely to be fairly accurate on average based on the independent results of Hegmann and Tavis [1998]. The mean age of our cohort was essentially the same as theirs:  $39.4 \pm 6.9$  and  $39.6 \pm 6.8$  years, respectively. Our cohort’s mean predicted  $\text{VO}_2$  max of  $32.8 \pm 7.1$  mL/kg · min is quite close to the directly measured aerobic capacity in their study:  $32.9 \pm 7.3$  mL/kg · min. In addition, the high morbidity group which had a higher proportion of the risk factors (increased age, smoking, lower % FVC, and adverse lipid profiles) as found by Hegmann and Tavis to predict  $\text{VO}_2$  max  $< 30$  mL/kg · min, had a predicted mean  $\text{VO}_2$  max of  $26 \pm 5$  mL/kg · min. Therefore, our regression equation might be helpful in identifying persons who are likely to have low aerobic capacities.

Regarding our CHD risk predictions, we used a model developed from a large cohort of white subjects drawn from a community in the western suburbs of Boston. Since our firefighters were predominantly white and from Massachusetts, the Framingham model should be generalizable to our cohort. Indeed, the mean CHD risks by age for our male firefighter cohort correspond closely with those found in the Wilson et al. [1998] Framingham study. A definite weakness, however, was the lack of complete CHD risk factor data (primarily cholesterol) for over half the cohort. The present article is based on the results of the first year of a 5 year longitudinal study. During the analysis of this data, we realized that cholesterol was not consistently being reported to the central data base. We have modified our reporting form to correct this deficiency. Therefore, we should have more complete information on CHD risk for the remainder of the study. Again, prospective follow-up of the cohort should reveal whether our predictions were accurate or not.

## CONCLUSIONS

Although criteria are needed to guide physicians when the presence of a single advanced or poorly controlled

condition may render a firefighter unfit for the safe performance of his/her essential job functions, poor fitness may be the result of a combination of several problems or risk factors. Based on examination of the data from our firefighter cohort, it seems that many of the least healthy or highest risk firefighters do not display extreme results on the tests usually considered in fitness determinations. Therefore, methods based on multiple risk factors or overall clinical assessments are superior means of identifying firefighters with poor health status and increased CHD risk when compared to selection methods based on single variables. The use of a multivariate CHD risk prediction formula appears to be the most accurate, cost-effective, evidence-based, and least arbitrary means of selecting firefighters for exercise stress testing. The age-based guideline suggested by NFPA is a sensitive selection tool, but has the disadvantage of selecting for stress testing a large proportion of firefighters who do not have a high CHD risk. It is still unclear, however, what threshold of increased CHD risk should trigger exercise stress testing. Such a threshold needs to be determined and validated in prospective studies of firefighter cohorts.

Our investigation also demonstrated that the group clinically rated below the 3<sup>rd</sup> percentile for health status (the high morbidity group) was a sub-group with high CHD risk, decreased pulmonary function, and low predicted aerobic capacity. The results suggest that firefighters clinically assessed to be in relatively poor health should undergo additional testing to further delineate their CHD risks and exercise tolerance. The minimum required aerobic capacity for firefighting remains unknown.

## ACKNOWLEDGMENTS

The authors would like to thank the members of Massachusetts District Hazardous Materials Response Teams, the Massachusetts Hazardous Materials Response Program, Chief Stephen Clendenin of the Commonwealth of Massachusetts Hazardous Materials Response Program, the staffs of Massachusetts Respiratory, Marlborough and Holyoke Hospitals, and Ms. Dianne Plantamura and Ms. Karen Encarnacion for their continuing support of this research. The authors would like to thank Mr. Stephen Foley and Dr. Richard Gerkin for their helpful comments.

## REFERENCES

Anderson KM, Wilson PW, Odell PM, Kannel WB. 1991. An updated coronary risk profile. A statement for health professionals. *Circulation* 83:356–362.

Boreham CA, Gamble RP, Wallace WF, Cran GW, Stevens AB. 1994. The health status of an ambulance service. *Occup Med (Oxf)* 44(3):137–40.

Davis PO, Dotson CO, Santa Maria DL. 1982. Relationship between simulated fire fighting tasks and physical performance measures. *Med Sci Sports* 14(1):65–71.

Gledhill N, Jamnik VK. 1992. Development and validation of a fitness screening protocol for firefighter applicants. *Can J Spt Sci* 17(3):199–206.

Hegmann KT, Tavis DR. 1998. Firefighters' fitness for duty: A study of the predictors for poor aerobic capacity. *J Occup Env Med* 40:1025.

Hines WW, Montgomery DC. 1980. *Probability and Statistics in Engineering and Management Science*. Critical values for the sign test (Appendix, Table X, p 621). New York: John Wiley and Sons.

Horowitz MR, Montgomery DL. 1993. Physiological profile of fire fighters compared to norms for the Canadian population. *Can J Spt Sci* 84(1):50–52.

Jette M, Sidney K. 1990. Fitness, performance and anthropometric characteristics of 19,185 Canadian forces personnel classified according to body mass index. *Military Med* 155(3):120–126.

Kales SN, Polyhronopoulos GN, Christiani DC. 1997. Medical surveillance of hazardous materials fire fighters: a two-year prospective study. *J Occup Env Med* 39:238–247.

Kales SN, Aldrich JM, Polyhronopoulos GN, Artzerounian D, Gassert T, Hu H, Kelsey K, Sweet C, Christiani DC. 1998. Fitness for duty evaluations in hazardous materials firefighters. *J Occup Env Med* 40:925–931.

Kales SN, Christiani DC. 1999. Fitness for duty evaluations of firefighters (response to letter to the editor). *J Occup Env Med* 41:214–215.

Kilbom A. 1980. Physical work capacity of firemen. *Scan J Work Environ Health* 6:48–57.

Lemon PWR, Hermiston RT. 1977. Physiological profile of professional fire fighters. *JOM* 19(5):337–340.

Morse L, Owen D, Becker CE. 1992. Firefighter's health and safety. In: Rom WN, editor. *Environmental and occupational medicine*. Boston: Little, Brown, p 1197–1204.

NFPA 1582. 1997 ed. *Standard on medical requirements for fire fighters*. Quincy, MA: National Fire Protection Association.

NFPA. 1998. *US firefighter fatalities, 1977–1996*. (slide show on NIOSH web page). Quincy, MA: National Fire Protection Association.

Occupational Safety and Health Administration. 1989. 29 CFR Part 1910, Hazardous waste operations and emergency response; Final Rule. Federal Register.

Pi-Sunyer FX. 1992. Obesity. In: Wyngaarden JB, Smith LH, Bennett JC, editors. *Cecil textbook of medicine*. Montreal: W.B. Saunders Co, p 1162–1170.

Samo DG, Bogucki S. 1999. Fitness for duty evaluations of firefighters (letter to the editor). *J Occup Env Med* 41:213–214.

Washburn AE, LeBlanc PR, Fahy RF. 1998. 1997 Firefighter fatalities. *NFPA* 92(4):50–62.

Wilson PWF, D'Agostino RB, Levey D, Belanger AM, Silbershatz H, Kannel WB. 1998. Prediction of coronary heart disease using risk factor categories. *Circulation* 97:1837–1847.