WORKING FOR A HEALTHY FUTURE

HISTORICAL RESEARCH REPORT

Research Report TM/88/08
1988

A review of current methods of fitness testing in the Mines Rescue Service and similar organisations

Love RG, Graveling RA

WORLD HEALTH ORGANISATION
COLLABORATING CENTRE
FOR OCCUPATIONAL HEALTH

A review of current methods of fitness testing in the Mines Rescue Service and similar organisations

Love RG, Graveling RA

INSTITUTE OF OCCUPATIONAL MEDICINE
Research Avenue North, Riccarton, Edinburgh, EH14 4AP
Tel: +44 (0)870 850 5131 Fax: +44 (0)870 850 5132
e-mail publications@iomhq.org.uk

Report No. TM/88/08
UDC 612.776.1 : 622.86

PHYSICAL FITNESS OF MINES RESCUE WORKERS

A review of current methods of fitness testing in the Mines Rescue Service and similar organisations

RG Love
and
RA Graveling

June 1988

Price: £20.00 (UK)
£25.00 (Overseas)

Report No. TM/88/08

INSTITUTE OF OCCUPATIONAL MEDICINE

PHYSICAL FITNESS OF MINES RESCUE WORKERS

A review of current methods of fitness testing in the Mines Rescue Service and similar organisations

by

RG Love and RA Graveling

Institute of Occupational Medicine
8 Roxburgh Place
Edinburgh
EH8 9SU

Tel: 031 667 5131
Telex: 9312100237=TD G

June 1988

This report is one of a series of Technical Memoranda (TM) published by the Institute of Occupational Medicine. Current and earlier lists of these reports, and of other Institute publications, are available from the Librarian/Information Officer at the address overleaf.

Fur further information about the Institute's facilities for research, service/consultancy and teaching please contact the Librarian/Information Officer in the first instance.

CONTENTS

INSTITUTE OF OCCUPATIONAL MEDICINE

PHYSICAL FITNESS OF MINES RESCUE WORKERS

A review of current methods of fitness testing in the Mines Rescue Service and similar organisations

by

RG Love and RA Graveling

SUMMARY

There is constant concern in the Mines Rescue Service that its members, whether full or part-time, should be of sufficient physical fitness to carry out their duties in an emergency without jeopardising the safety of their colleagues in the rescue team. The current position with regard to fitness testing procedures in the rescue service is described and compared with the situation in the Fire Service and other organisations where physical fitness is considered to be a prerequisite of membership. The potential disadvantages of using a fitness index originally derived from the near-maximal Harvard Pack Test and applying it to a submaximal treadmill test are discussed.

The available submaximal tests are reviewed for their usefulness to predict physical fitness and the different modes of testing (treadmill, bicycle ergometer, etc.) and test protocols are described. The inherent inter and intra-test variability of such tests, the use of submaximal or maximal tests and their relative advantages and disadvantages are also discussed with a view to recommending improvements in the test currently used in the Mines Rescue Service. In particular, it is suggested that a fundamental reappraisal is made of the way in which the test is used, in order to avoid too much reliance on the rigid pass/fail point given by the derived Harvard Pack Index.

1. INTRODUCTION

Mines Rescue Service of British Coal is responsible for the rescue of underground personnel, the containment and extinction of underground fires and activities associated with other emergencies in coal mines or similar environments, where their expertise is required, e.g. sewers, tunnels, etc. As in other rescue services the job requires a team of highly selected men with an excellent standard of physical and mental fitness. Not only could the life of a miner depend on the fitness of a rescue man, but the safety of other rescue men in a team could be jeopardised by the presence of one or more members having less than the required physical fitness.

The Mines Rescue Service consists of fulltime men at central rescue stations and many more part-timers, mineworkers at local collieries, who shoulder the major burden of rescue work. Historically this has been advantageous in that the generally strenuous nature of mining work has meant that the majority of rescue workers are kept sufficiently fit. Permanent corps men do not have this advantage and, with increasing mechanisation in mining tasks, part-time men are also not undertaking as much physically demanding work. There is therefore an increasing need for a form of fitness testing to ensure that all rescue brigadesmen are fully fit for their tasks in a colliery emergency. Both full and part-timers have to leave active service at the age of 45 years.

This note aims to review the current situation with regard to fitness testing in the Mines Rescue Service and in other services where a selected fit group of men (and women) may be required to undertake strenuous activities, including rescue work, often while wearing self-contained breathing apparatus, e.g. firefighters, deep sea divers, military service personnel.

Currently available methods of submaximal exercise testing are compared and reviewed. These and possible amendments to existing procedures in the Mines Rescue Service are discussed.

2. CURRENT TESTING PROCEDURES

2.1 Mines Rescue Service

Following trials to establish the mines rescue service with 'safe working times' in hot and humid environments it was recommended that an exercise tolerance test (and a heat tolerance test) should be included during the selection process for rescue brigadesmen (Lind and McNicol, 1968). The statutory annual medical examination of all mines rescue personnel and potential recruits now demands that a basic physical fitness test is passed each year (Spencer, 1972). In the late 1960s the Harvard Pack Test was introduced as the test, which best assessed physical fitness, although it requires determination as well as stamina to complete. In the test the man carries one third of his body weight on his back and, with the aid of handrails, steps on and off a 40cm step 30 times per minute for up to 5 minutes. It was based on the test, devised by Johnson *et al* (1942) as a rapid, simple method of assessing fitness for strenuous work.

The test was originally designed, so that one third of all subjects stopped from exhaustion within 5 minutes. It allows a fitness index to be calculated from the total duration of the exercise and the sum of three recovery pulse counts. This index can be used following any appropriately exhausting exercise and in this case the Harvard Pack Index (HPI)

$$= \frac{\text{Duration of stepping exercise in seconds x 100}}{\text{2 x sum of pulse counts from } 1\text{–}1\tfrac{1}{2},\ 2\text{–}2\tfrac{1}{2},\ 4\text{–}4\tfrac{1}{2} \text{ minutes of recovery}}$$

A good result is any index greater than 75, and 100 is very good. A pass value of 75 was introduced for the Mines Rescue Service.

However, although many thousands of such tests were undertaken by the Rescue Service, the death of a rescue brigadesman in South Wales in 1972 thirty minutes after completion of the Harvard Pack Test led to a reassessment of the suitability of such a severe test and investigation of a possible alternative. As a result a study was undertaken by the NCB Medical Service, S. Wales Area, and the MRC Pneumoconiosis Unit, to determine if a series of sub-maximal exercise tests could be used to replace the Harvard Pack Test as a guide to the cardio-pulmonary response to exercise (Cotes *et al*, 1979).

The results of this study led to the introduction of a submaximal treadmill test for six minutes duration with continuous electocardiographic monitoring, from which the last minute cardiac frequency (LMCF) was calculated (Cotes *et al*, 1979). The authors used their results to predict a Harvard Pack Index (HPI_{Pr}). Correlation of the measured HPI with anthropometric and physiological variables was also investigated. The HPI could be predicted from combinations of these other indices with coefficients of variation of about 10% and provided the basis for the test index currently used by the Mines Rescue Service.

Cotes *et al* (1979) investigated a number of relationships between the HPI and other measured variables and concluded that the most practicable was based on age and the last minute cardiac frequency (LMCF), the eleventh minute in their original test. This relationship was expressed as follows:

HPI_{Pr} = 147.1 - 0.215 age - 0.39 LMCF (Coefficient of variation = 10.4%)

HPI_{Pr}, when plotted against the measured HPI, had a wide range of values at any given HPI, particularly at higher levels. For example, at an HPI_{Pr} of 96 (quite a high score) there were measured values ranging from 78 to 132. More imporantly for the individual rescue man 27 men failed the Harvard Pack Test (score <75) but did not also fail the treadmill test. On the other hand only four men failed the treadmill test but not the pack test. These, possibly chance, effects indicate the limitations to any indirect procedure for describing the performance of individuals, as the authors of this study pointed out.

Nevertheless the six minute treadmill test (speed 5 kph, incline 7.5^{o}) is now the standard method for assessment of the cardiovascular fitness of rescue brigadesmen in the UK coal industry. This has led to a number of possible disadvantages being reported within the rescue and medical services.

1. Whereas the pack test stretched many men to their physical limit, the treadmill test is a sub-maximal test for all men, who, in some cases, do not feel that they have been fully tested. This has been partly overcome in at least one British Coal area, in which the test is extended by two minutes at a higher treadmill speed, individually programmed to increase each man's heart rate to about 80% predicted maximum. An additional index is derived, based on a ratio of the measured cardiac frequency (7th - 8th minute) to that previously predicted from the instantaneous 5th minute rate, a lower ratio intended to indicate better fitness.

2. Because of the implications of the medical examination on future or continued service as a rescue brigadesman, some men may feel anxious before and during the test. This may elevate the heart rate during the period of exercise particularly for those of good fitness, whose heart rates are not raised substantially by the test. As a result such individuals may fail to achieve the required standard of fitness index, which is equivalent to 75 on conversion to the HPI.

3. The treadmill test only provides an assessment of aerobic fitness during submaximal leg exercise and provides no indication of muscular strength of the arms and shoulder girdle. The Harvard Pack Test utilises the arms, legs, back and many accessory muscles and may therefore provide a better overall indication of fitness for mines rescue work.

4. In the case of new recruits unfamiliarity with walking on a moving belt at the first attempt may elevate cardiac and pulmonary functions compared with subsequent occasions (Cotes, 1979) and adequate practice may not always be allowed. However, the same could also be said of the pack test, which is often more difficult to undertake correctly, i.e. by lifting the body through its full centre of gravity at each step.

It has already been pointed out that the increasing mechanisation and technological development in coalmining has led to reduced opportunities for part-time rescue men to maintain physical fitness. This problem has been recognised in other developed countries and some attempts have been made to introduce a structured fitness programme into the rescue worker's training. In West Germany the results of a pilot study of the effectiveness of regular circuit and exercise bicycle training have been very encouraging (Schulte, 1983). Attempts have been made to provide exercise equipment at rescue stations in this country but regular, structured training sessions are necessary for optimum benefits to be obtained.

2.2 Fire Service

In many ways the requirements of the Fire Service are similar to those of the Mines Rescue Service, but the fire man or woman is expected to tackle a wider range of unknown conditions at a moment's notice and at more frequent intervals. The firefighter is also trained in the use of breathing apparatus (currently open circuit compressed air sets) and needs to have a good physical fitness in order to qualify for the job.

Although at present no formal test of physical fitness is applied either to recruits to the Fire Service or to those undergoing re-examination, physical standards are detailed in the Fire Services (Appointments and Promotion) Regulations 1965. These state that an applicant shall have satisfied a medical practitioner selected by the fire authority, that he is fit to undertake fire fighting duties.

Recommendations of the Godber Report of the Committee to Review the Medical Standards for the Fire Service in 1968 stated "operational firemen may be subjected to considerable physical strain and have at times to exert extreme physical effort, sometimes prolonged, under very adverse conditions of atmosphere and heat. We consider ... that his fitness, particularly his cardio-respiratory function, should be subject to periodic review."

However, the only stipulation was that after the age of 40, firemen should be examined every three years. The form of these repeat examinations was not defined but it was suggested that the general guidance given to medical practitioners and the report form used for the initial examination might be the same. The guidance for medical standards states that the physical capacity should be such that the fireman should be fit for manual work, including lifting and climbing; that upper limb muscle power be average, but able to cope with heavy manual work, and that he should be capable of running, climbing ladders, jumping, crawling and performing all kinds of manual labour under conditions expected at a fire. No tests are specified to quantify these requirements and only a resting electrocardiogram is taken routinely.

Unlike most other occupations firefighters (and, to some extent, mines rescue workers) are not able to control the pace of their work to suit their capacity but must respond to the ever-changing needs of the emergency. Physical activity in a simulated fire fighting environment has been found to require 97% of V_{02} max in some cases, which does not leave a great margin for safety (Davis and Dotson, 1978). This is equivalent to an energy expenditure twelve times greater than resting. Routine fire fighting tasks, such as working with hoses, have been shown to require 60-80% of aerobic capacity (Lemon and Hermiston, 1977). Much of this energy requirement is extra expenditure necessitated by the wearing of heavy protective clothing and breathing apparatus (about one-third body weight) which can increase oxygen consumption by 20 - 30% during simple treadmill walking (Davis and Santa Maria, 1975; Louhevaara *et al*, 1984).

Firefighters therefore need to be physically very fit just to overcome the disadvantage of working with at least 30% less efficiency. Yet previous studies have shown that the average maximal aerobic capacity of firemen in several countries, including the U.K., may differ little from that of sedentary men (Brown *et al*, 1982; Kilbom, 1980; Kuorinka and Korhonen, 1981; Louhevaara *et al*, 1985). Consequently the need for organised and structured fitness training programmes has been recognised by several fire services around the world and in some cases details have been published (Davis *et al*, 1982; Bahrke, 1982; Brown *et al*, 1982 and Cady *et al*, 1985).

Although firefighters may not have greater than average physical fitness, they have to pass a stringent medical examination, in order to exclude a wide range of diseases. It has been suggested, however, that the high stress imposed on firefighters by their job might be associated with increased morbidity and mortality from cardiovascular diseases. The potential danger from stress in firemen with undiagnosed myocardial problems has also been emphasised (Manning and Griggs, 1983). An earlier study of the mortality of firemen (Mastromatteo, 1959), which identified an increased mortality from cardiovascular and renal diseases (SMR = 135), has not been confirmed in subsequent studies (e.g. Musk *et al*, 1978; Eliopulos *et al*, 1984).

As a result of recent discussions, recommendations are being put forward by a working party on appointment provisions to the Fire Service, suggesting that a physical fitness test be applied to recruits and those undergoing periodic reexamination. Although different possibilities have been considered, it is likely that in practice some form of step test will be recommended for its simplicity and ease of operation. These recommendations follow an extensive study of the physical fitness of firemen by Chelsea College, University of London (personal communication), which established the required level of physical fitness for firemen.

2.3 Commercial diving

With the advent of the North Sea oil and gas industries the demand for deep-sea divers with commercial experience has increased rapidly and many more divers have been recruited who have not had naval diving experience. As in the rescue services, divers often work in teams and it is therefore imperative that the poor fitness of one member should not put the others at any unnecessary risk. However, the nature of the divers' work is likely to be more periodic, especially for those employed in North Sea operations.

Fitness to dive in British waters is covered by the Diving Operations at Work Regulations (1981) and doctors approved by the Health and Safety Executive conduct examinations and issue certificates of fitness to dive. Professional divers are normally required to be under 40 years of age, although this need not be an excluding factor, and to have had a medical examination specifically for fitness to dive in the previous year. They must show evidence of good standard of physical fitness, the minimum requirement being the absence of obesity. As stated in the authoritative 'Physiology and Medicine of Diving' (Elliott and Davis, 1982): "In spite of the fact that the limiting factor in such circumstances is respiratory, an exercise tolerance test can be used to demonstrate the cardiovascular reserves of an individual who may have to work maximally in an underwater emergency."

As well as excluding relevant respiratory, cardiovascular, gastrointestinal, neural or psychiatric conditions, which might predispose to medical emergencies underwater, the Regulations (Health and Safety Executive, 1981) state that an exercise tolerance test should be carried out. In practice the examining doctor assesses the fitness by careful examination and history taking e.g. current active recreations, sport, training, etc. If the doctor is in any doubt over the man's physical condition, a fitness test can be recommended. This is suggested to be the Army Physical Fitness Test (see later) but it can take the form of a submaximal treadmill test (used at one centre in Scotland) or a step test (used at a diving centre in England). For these reasons it is likely that a wide range of fitness will be observed among commercial divers.

In this context it is of interest to note the report of a recent study of 148 experienced and active North Sea divers (Thompson *et al*, 1984). Based on a submaximal treadmill exercise these authors observed an average maximal oxygen uptake (V_{O2}max) of 46 $\pm$ 9 ml O_2 kg^{-1} min^{-1} for these divers. This compared with 41 $\pm$ 8 ml kg^{-1} min^{-1} for 130 normal, healthy but sedentary men and 60$\pm$ 9ml kg^{-1} min^{-1} for recreational runners. Although the V_{O2}max was consistent with the average for men of the same age, i.e. 40–51 ml kg^{-1} min^{-1} (Astrand, 1960), the authors stated that such divers should be fitter than this for their particular work and suggested a minimum MAP (maximal aerobic power) of 50 ml kg^{-1} min^{-1}.

Cotes and Reed (1984) commented that there was no known evidence to support such a minimum requirement, as much of the diver's work is of a sedentary or static nature, which does not promote cardiovascular fitness. These and other (Crosbie, 1984) authors emphasised the value of experience rather than an arbitrary level of MAP. A recommendation to provide a suitable standardised exercise test in the annual medical, instead of an optional test, has been made as a means of providing the evidence of the need for the proposed standard by Cotes and Reed (1984). These authors also observed that only 50% of mines rescue brigadesmen, who have to pass an annual fitness test to remain in the rescue service, maintained their aerobic power at this proposed level (Robertshaw *et al*, 1984).
It is likely that the divers who have to undergo a fitness test are less fit than those who, in the opinion of the examining doctor, are not required to submit to such a test. Until exercise tolerance tests are a requirement in a medical examination the question of the need for minimal fitness standards cannot be satisfactorily resolved.

2.4 Military personnel

The army and other military services have a need for fit, highly trained personnel who will work together as a team. All British army recruits are expected to undergo a basic fitness test (BFT), which is used as the main indication of personal fitness and the fighting unit's physical fitness for its role. BFT1 is taken by soldiers under 40 at least twice yearly. It consists of two parts: part 1, a run and walk of 2.4km in 15 minutes in a squad, followed immediately by part 2, a best effort run of 2.4km in under 11½, 12 or 13 minutes, depending on age. This test is performed in boots, anklets, trousers and vest. Studies carried out in Canada (Myles *et al*, 1980) and the UK (Myles and Toft, 1982), have shown that the 2.4km running time correlates very well with directly measured V_{O_2max} and is therefore considered to represent a valid measure of cardio-respiratory fitness.

A modified Harvard Step Test (the Army Physical Fitness Test referred to in the previous section) is also employed in the British Army by physical training instructors as a guide to planning individual training for soldiers. It consists of stepping on and off a 43 cm step, 30 times per minute for 5 minutes. The sum of the three ½ minute pulse counts, taken at 1, 2 and 3 minutes after the exercise has stopped, is converted to a test score, which is used to grade the individual into high or satisfactory grades according to the same age ranges as in the BFT (17–29, 30–34 and 35–39 years). The procedure for this and other tests of fitness are described in the Army's "Fit to Fight" pamphlet.

3.SUBMAXIMAL EXERCISE TESTING FOR CARDIO-RESPIRATORY FITNESS

3.1 Introduction

Many exercise tests have been described in the literature. Some, such as those described by Bruce *et al* (1963) and Erkan (1977), have been devised to screen for symptoms of coronary heart disease (frequently known as a 'stress test'). Others, e.g., Foehr (1974), are intended to quantify reactivity to work in the heat. A third group, however, has been developed primarily as a means of assessing what is loosely termed 'fitness', but which is generally intended to evaluate cardio-respiratory function during physical work (physical work capacity). It is this latter group which is examined in this overview of the literature.

3.2 Overview of submaximal tests

Numerous submaximal exercise tests have been developed over a period of many years and it would be too big a task to detail them all, together with all the slight variations which have been reported in other papers. Although the purposes for which these tests were developed differ from test to test, many share a common aim of providing some prediction of maximum physical work capacity (usually equated with maximum oxygen uptake) without subjecting individuals to the high levels of stress usually associated with a direct assessment of maximum capability.

Reports as to the predictive ability of these tests vary considerably. For example, Burke (1976) reported various studies where comparisons had been made between measured and predicted maximum oxygen uptake using the Astrand-Ryhming prediction. The correlation coefficients obtained ranged from r = -0.43 to r = +0.75 (the sign depending upon the units used). This variability may reflect differences in the fitness and motivation of the groups tested (Graveling, 1978), particularly in the satisfactory direct determination of maximum oxygen uptake. A survey by the authors of a sample of comparison studies using a variety of step, cycle and treadmill tests revealed correlation coefficients ranging from 0.21 to 0.98. At least as important was the observation that the coefficients of variation (c.v.), which give an indication of the accuracy of any estimate, ranged from 5% to 15%. Two particular tests stood out from this survey as having a high linear correlation with maximum oxygen uptake and a low coefficient of variation (either reported in the test or calculated from data given).

1. Van der Merwe *et al* (1980) reported the results from a step test, in which the bench height was adjusted according to body weight to give a constant work load. Tests were carried out on a variety of groups, including military recruits and various groups of sportsmen, all with mean ages of less than 26 years. Maximum oxygen uptake was determined directly on two sub-groups of 11 military recruits and 4 canoeists. Despite a significant difference in maximum oxygen uptake between these two subgroups, the data were pooled for multiple regression analyses. The best regression was one involving body weight, the heart rate at the end of exercise (eight minutes) and the heart rate after one minute of recovery. This yielded a correlation coefficient of 0.94 with a standard error of estimate of 0.173 litres per minute. A further regression correlation was reported, apparently involving the same variables, with a slightly lower multiple correlation coefficient (r = 0.91) but also a lower error estimate (= 0.144 litres per minute). When

compared to the mean maximum oxygen uptake reported for the 11 recruits in the subgroup, these yield coefficients of variation of 5.0% and 4.2% respectively. It is interesting to note that, although body weight was taken into account in setting the work task, its inclusion in the regression produced a marked improvement in correlation coefficient and the accompanying standard error.

2. Similar coefficients of variation can be calculated from the results reported by Legge and Banister (1986). These authors redrew a nomogram for the prediction of maximum oxygen uptake along similar lines to that in the original Astrand-Rhyming nomogram (Astrand and Rodahl, 1977). However, a major difference was the replacement of the absolute heart rate in the original nomogram with a delta heart rate value, representing the elevation in heart rate above that from an initial 'zero-load' work task (pedalling a cycle ergometer with no added resistance). The intention through this was to reduce the influence in the heart rate term of factors such as emotion and ambient conditions. A comparison of predicted versus observed maximum oxygen uptakes using the new nomogram and the original Astrand-Ryhming nomogram yielded linear correlations of 0.98 and 0.80 respectively, a statisitically significant improvement ($p = 0.05$). The standard error of the estimate associated with the new nomogram from this comparison was 0.17. When compared to an approximate average maximum oxygen uptake, derived from the graph, of 3.5 litres per minute, this yields a coefficient of variation of 4.9%. The low c.v. values from these two techniques are particularly interesting as they are derived from heart rate values. Shephard (1966) has suggested that improved predictions are obtained by using measurement of oxygen uptake at submaximal levels, rather than heart rate, in any predictive equation.

Davies (1968) reported the results of an extensive evaluation of the relationship between oxygen uptake and heart rate with particular reference to the prediction of maximum oxygen uptake. Using a series of established predictive nomograms including that of Astrand and Ryhming, Davies found predicted maxima to be consistently lower than those measured. The author attributed this poor predictive ability of the submaximal exercise tests to the asymptotic relationship between cardiac frequency and oxygen consumption. Perhaps more important, however, than this systematic error was the large non-systematic error associated with the techniques. Thus, Davies calculated that, using a single measurement of heart rate between 120 and 150 beats per minute, the 95% probability limits around the predicted value would be within 1.22 litres per minute of the directly observed value. The author concluded that, if an accuracy of maximum oxygen uptake greater than ±15% was required, then there was no alternative but to measure it directly. Davies made a few recommendations for improving the predictive accuracy of submaximal techniques based upon his findings, including using more than one cardiac frequency value or constructing an oxygen uptake/cardiac frequency curve for each individual. Although the author criticised these approaches for still underestimating the measured maximum, perhaps their main benefit lies in the notably smaller standard deviations of the prediction. A further suggestion was the use of a relatively high submaximal workload, a suggestion which is supported by results reported by Legge and Banister (1986).

The relationship between the heart rate response to exercise and maximum work ability, questioned by Shephard (1966) and others, was studied in detail by Brooke and Hamley (1972). They reported three main components to the response, an initial anticipatory response, a linear phase, and a final negative quadratic phase. The anticipatory response would be accounted for in the approach of Legge and Banister in using a delta heart rate value. Brooke and Hamley showed that the slope of the linear phase did not correlate with maximum work ability. As this forms the basis of many predictive approaches in which submaximal heart rates are extrapolated to maximal values, it is not therefore surprising that they do not produce reliable predictive indices. The authors reported that the only phase which correlated significantly with maximum work capacity was the final curvilinear phase with a high curvature being associated with poor work ability.

3.3 Comparison of different modes of testing

An early comparison of different fitness tests was reported by Eichna *et al* (1944) on behalf of the US Army. The tests varied in complexity from the Harvard step test (stepping up and down on a 20 inch platform once every 2 seconds for 5 minutes) to the Army Ground Forces (AGF) test consisting of a battery of six tests: number of push-ups, time to run 300 yards, the number of Burpees in 30 seconds, time to run 75 yards carrying a man of equal weight, time to accomplish (sic) 70 yards creeping, crawling, jumping and running and a march of 4 miles in 50 minutes. The authors concluded that men tested attained differing and variable ratings by the different tests. For example, on a three way rating of poor, average and good, of the 125 men tested on two tests, 52 had the same rating on both whereas 73 differed. The authors also concluded that:- 'an alert, interested officer who has worked with and knows his men is capable of giving a better evaluation of the fitness (both physical and mental) of his men than any fitness test yet devised', although no objective evidence for this was reported. As a corollary to this, it was suggested that fitness tests seemed to be of most value as an adjunct to arriving at a final concept of fitness for work rather than the definitive determinant.

Shephard (1966) reported a comparison of a step test, a bicycle ergometer test and a treadmill test in the assessment of cardio-respiratory fitness. He compared sub-maximal predictions of maximal aerobic capacity (using the seemingly ubiquitous Astrand-Ryhming nomogram) with direct measurements. There was little to choose between the three sets of results. The mean treadmill prediction (reported as a difference between measured and predicted of +0.03 litres) was closest to the measured value but the variation about the mean was the highest (14.9%). As referred to earlier, a more consistent prediction was obtained in all three cases by using oxygen consumption and pulse rate rather than work level and pulse rate, with coefficients of variation some 30% lower. Shephard concluded that there was no benefit to be gained from using a treadmill or bicycle rather than a step test from the point of the accuracy of predictions but re-emphasised that, whichever method was used, work capacity could be more consistently predicted if oxygen consumption values were obtained.

Treadmill walking, cycling and stepping were also examined by Cotes *et al* (1967). In this short paper, the authors reported that greater precision of submaximal exercise ventilation and cardiac frequency could be obtained by determinations at a fixed level of oxygen uptake rather than a fixed work level, a similar approach to that embodied in the PWC150 and PWC170 tests (Physical Work Capacity) where the work capacity is measured at the specified heart rates of 150 or 170 beats min^{-1}.

Shephard (1966) used a step test to derive maximum oxygen uptake values for his comparisons but this did not yield any apparent preference for submaximal predictions using stepping tests. In contrast, Burke (1976), in an extensive comparison of tests of physical working capacity, used a treadmill based determination of maximum oxygen uptake as a benchmark. Seven laboratory tests and seven field tests were compared for predictive ability. The laboratory tests involved treadmill tests such as the Cureton all-out run, (10mph, 8.6% grade) cycle ergometer tests such as the PWC170 and step tests such as the Margaria test (time sprinting up stairs). The field tests ranged from the 10 yard dash to the Cooper 12 minute run. In this case, the best correlations with measured maximum oxygen uptake were obtained from the longer running tests, particularly the 12 minute run (r = 0.90), and the treadmill tests (r = 0.69 - 0.77) i.e. those tests which were most comparable to the maximum test. Such tests also tended to have smaller standard errors of the estimates. It was, however, acknowledged that the field tests, requiring maximal effort, were dependent on a high level of motivation, which, although obtained in the group tested, may not always be readily achieved. Of the tests described as submaximal, the Balke test, involving treadmill walking at progressively increasing gradients, was recommended as the best of those examined (r = 0.77). It was also concluded that tests of longer duration provided better measures of aerobic power.

Robertshaw *et al* (1984) reported a comparison of four different tests against directly measured maximum oxygen uptake from a treadmill running test. The tests were the Harvard pack test, treadmill walking, cycling and a step test. Although the four tests correlated reasonably well with each other, with correlations ranging from -0.52 to 0.80, correlations with maximum oxygen uptake were generally poor (r = 0.21 to 0.37). The authors concluded by drawing attention to these poor correlations suggesting that the submaximal indices were of little predictive value. This was seen as being consistent with other reports which had demonstrated the limited value of cardiac frequency for predicting maximal oxygen uptake.

In one interesting comparative study of different modes of exercise testing Leblanc *et al* (1981) compared their degree of agreement rather than their predictive ability. They examined different modes of testing to determine to what extent the tests could be regarded as examining a common attribute of submaximal working capacity. The authors compared leg cycling (both supine and sitting), arm cranking standing, treadmill walking and bench stepping at submaximal levels of work (40-55% of maximum). They examined the shared variance between the different tests for different physiological parameters. As might be expected, the shared variances differed between test modes and physiological parameters. The highest multiple correlations (and hence shared variance) were generally obtained between supine cycling and the other forms of exercise and for the submaximal heart rate values. The bivariate shared variances between treadmill walking and bench stepping ranged from 52% for heart rate and oxygen consumption in litres per minute to 74% for minute ventilation. However, when body weight was taken into account, by expressing oxygen consumption values per kilogram of body maximum oxygen uptakes obtained, the highest values being obtained from the Taylor protocol. Furthermore, although not statistically significant, the Taylor protocol had lower coefficients of variation averaged over the fifteen subjects with three repeats of each test. The average coefficients reported for these maximal tests ranged from 4.1–5.8%, noticeably lower than the values of 10–15% typically reported for predictive submaximal tests.

The concept of an intermittent test protocol was examined in a more systematic fashion by Fardy and Hellerstein (1978), who examined a protocol with stepped increases in gradient and speed, with and without a rest period between each stage. Predicted maximum oxygen uptake was 6% higher in the intermittent test protocol than in the continuous protocol, although this difference just failed to attain statistical significance ($0.05<p<0.10$). Additionally however, all subjects perceived the intermittent test as less strenuous and were able to complete at least one additional stage. Physiological measures were indeed generally lower for the intermittent test than for the equivalent stage in the continuous test.

In an interesting variation on the concept of gradually increasing workload, Nagle *et al* (1965) studied a step test where the step height was increased at intervals during the test. Two tests were reported: the first was a 30 steps per minute test where the platform height started at 2.0cm and was increased by 2cm every minute. In the second, 24 step test, the platform height started at 3.3cm but was only raised by 1.7cm each minute of the test. Unfortunately, although the paper demonstrated the variation in oxygen requirement for the two tests, no attempt was made to assess their relative suitability for determining maximum oxygen consumption or to compare them with a conventional step test.

4. DISCUSSION

It is apparent from the papers reviewed that submaximal exercise testing does not provide an accurate measure of physical work capacity as indicated by maximum oxygen uptake. Predictive approaches intended for a broadly defined population appear generally to have an inherent inaccuracy of approximately $\pm$10% when compared with direct measurements. Personal experience (e.g. Graveling, 1978) has shown that, with non-athletes, both predicted and measured oxygen uptake can vary considerably between tests and that, even with established international athletes and sportsmen, the testing environment and procedure can have a considerable effect on the results obtained. It is this inter- and intra-test variability which gives most cause for concern. Inability to obtain an accurate indication of working capacity can be tolerated if the inaccuracy is consistent, but obtaining an unreliable indication, either in comparisons between individuals or on the same individual on different occasions, can clearly cause problems when the test forms the basis of a screening programme.

The development work for the treadmill test currently used by British Coal took comparability with the Harvard Pack Index (HPI) rather than oxygen consumption as its criterion (although the HPI was originally related to oxygen consumption) and much the same criticisms apply. In particular, the coefficient of variation of 10.4% associated with the HPI predictive equation (Cotes *et al*, 1979) implies that a predictive value of 75 has a reasonable chance of indicating a true score between 68 and 82 (with approximately one in three true scores falling outside that range).

The literature referred to earlier has indicated ways of improving the predictive ability of various submaximal tests. Most of the authors have concentrated on improving their accuracy, although some may equally apply to improving consistency. Amongst these recommendations has been the use of direct measurement of oxygen consumption or heart rate changes rather than the absolute level of heart rate. Both of these parameters have the potential benefit of removing or reducing the influence of extraneous psychological or environmental factors on the physiological responses and are worthy of further evaluation.

The argument has also arisen as to whether a maximal (or near maximal) test or a truly submaximal test of physical fitness should be used to assess exercise tolerance among those likely to be exposed to strenuous working conditions in the rescue and other services. In the mines rescue and fire services men and, in a few cases, women experience short periods of very intense activity (carrying hoses, stretchers, stopping materials, etc) and longer periods of sustained moderate activity (damping down fires, walking or crawling in tunnels, etc.). These activites may be undertaken in very hot and humid conditions, invariably while a self-contained breathing apparatus is worn and, in the case of firefighters, in heavy duty protective or impermeable clothing.

In combination these stressors can impose a severe load on the individual. It is clearly an advantage that, for a given task or work load, the person with the greater physical working capacity is able to operate well within his or her aerobic capacity, thus reducing the chance of the onset of fatigue during a period of duty. The case against some form of test to determine an agreed minimum level of fitness for such work is probably unarguable but the type of test is still disputed.

Some of the objections to the submaximal test have been raised earlier but this does not imply that one should go to the other extreme and load a potential recruit with heavy clothing and apparatus and then exercise him to exhaustion (e.g. Raven *et al*, 1977) possibly in hot and humid conditions. Besides the ethical considerations (even with medical cover and electrocardiographic monitoring), such a test would probably serve no useful purpose and might even have the reverse effect to that intended. Experience of heat and humidity and wearing breathing apparatus during strenuous activity can be simulated in training periods in realistic settings. A physical fitness, exercise tolerance, cardiovascular fitness, aerobic capacity test, however it is described, is primarily intended to quantify under controlled and repeatable conditions an index of physical working capacity, although not specifically muscular strength.

For those in the emergency services, who are usually highly motivated, competitive individuals, a test which presents an easily recognisable, physical challenge is probably desirable. it is also highly desirable that any underlying cardiovascular, orthopaedic or other conditions, liable to render an individual unsuitable for such a job, are detected in a controlled environment rather than occurring in the course of a real emergency. An exercise which utilises large muscle groups (the legs) is to be preferred and, if continuous electrocardiographic monitoring is to be used, this also reduces interference with the ECG signal from extraneous muscular contractions in the upper body.

An upper limit to the severity of the exercise can be set by increasing the workload to a point at which 80% or even 90% of the predicted maximal heart rate is reached. The participants should feel that they have been given a reasonably strenuous test (of the same relative severity for each person) and one or more indices can be derived from the recorded variables (see section describing tests) to allow the observer to obtain a satisfactory fitness score.

A possibility worth considering is to conduct the fitness test in the manner that Johnson *et al*'s (1942) original Harvard test was intended, i.e. a treadmill run at a speed of 7mph on a gradient of 8.6%. A slower speed (to allow walking) and steeper gradient could be used to give an equivalent workload, in order to facilitate ECG recording.

However, there would appear to be no 'ideal' test medium. Step testing, treadmill testing and bicycle testing all have their exponents and also have their inherent advantages and disadvantages. For example, problems have been reported with differences in calibration of cycle ergometers (Robertshaw *et al*, 1984). Equally, however, the American Heart Association deemed it necessary to issue detailed specifications for calibrating both the speed and the elevation of treadmills for exercise testing (Hellerstein, 1979).

Perhaps more fundamental to the 'fine tuning' of a particular exercise test is the way in which the test is regarded. As reported by Eichna and coworkers as early as 1944, fitness tests should be used as an adjunct to determining 'fitness for work' rather than the absolute arbiter. While the technology of exercise testing has improved dramatically since that opinion was expressed, the reasoning behind it, as outlined above, would still seem to be equally valid.

5. CONCLUSION

Submaximal predictive tests of work capacity appear to be inherently inaccurate and unreliable for use in isolation for predicting working capacity (which is of course only one part of overall 'fitness for work'). The variation in predictive ability of such tests implies that there is a strong possibility when using such a test in isolation of either failing an individual unnecessarily or, perhaps more seriously for the organisation, passing someone in error.

Improvements to the test currently used in the Mines Rescue Service and to the way in which it is administered could reduce the effect of this inaccuracy but these are unlikely to remove it completely. What would appear to be needed is a more fundamental reappraisal of the way in which the test is used, with a move away from a dogmatic view of '74 predicted = fail, 76 predicted = pass' to reflect this uncertainty.

ACKNOWLEDGEMENTS

This review was undertaken with financial support from British Coal, Medical Service, to whom we are most grateful.

REFERENCES

AMERICAN HEART ASSOCIATION SUBCOMMITTEE ON REHABILITATION, TARGET ACTIVITY GROUP, HELLERSTEIN HK. (Chairman). Specifications for exercise testing equipment. Journal of Occupational Medicine 1979; 21: 828-832

ASTRAND I. Aerobic work capacity in men and women with special reference to age. Acta Physiologica Scandinavica 1960; 49(suppl.169):

ASTRAND PO, RODAHL K. Textbook of work physiology. Physiological bases of exercise. 2nd ed. New York: McGraw-Hill, 1977.

BAHRKE MS. Voluntary and mandatory fitness programs for fire fighters. Physician and Sportsmedicine 1982; 10: 126-132

BINKHORST RA, VAN LEEUWEN P. A rapid method for the determination of aerobic capacity. International Zeitschrift fur angewandte Physiologie 1963; 19: 459-467

BROOKE JD, HAMLEY EJ. The heart-rate - physical work curve analysis for the prediction of exhausting work ability. Medicine and Science in Sports 1972; 4: 23-26

BROWN A, COTES JE, MORTIMORE IL, REED JW. An exercise training programme for firemen. Ergonomics 1982; 25: 793-800

BRUCE RA, BLACKMON JR, JONES JW, STRAIT G. Exercising testing in adult normal subjects and cardiac patients. Pediatrics 1963; 32: 742-756

BUCHFUHRER MJ, HANSEN JE, ROBINSON TE, SUE DY, WASSERMAN K, WHIPP BJ. Optimizing the exercise protocol for cardiopulmonary assessment. Journal of Applied Physiology 1983; 55: 1558-1564

BURKE EJ. Validity of selected laboratory and field tests of physical working capacity. Research Quarterly 1976; 47: 95-104

CADY LD, THOMAS PC, KARWASKY RJ. Program for increasing health and physical fitness of fire fighters. Journal of Occupational Medicine 1985; 27: 110-114

COTES JE. Lung function: assessment and application in medicine. 4th ed. Oxford: Blackwell Scientific Publications, 1979.

COTES JE, DAVIES CTM, JOHN C. Relationship of oxygen uptake, ventilation and cardiac frequency to body weight during different forms of standardized exercise. Journal of Physiology 1967; 192: 29P-30P

COTES JE, DICKEN C, EVANS DL, JOHNSON GR, KALINOWSKA EB, MACINTYRE IM, SAUNDERS MJ. Prediction of Harvard pack index from the result of an 11 min progressive exercise test and anthropometric measurements in coalminers. Ergonomics 1979; 22: 1353-1361

COTES JE, REED JW. North Sea divers are no fitter than sedentary men. Letter. Lancet 1984; ii: 348-349

CROSBIE WA. Fitness of North Sea divers. Letter. Lancet 1984; ii: 471-472

DAVIES CTM. Limitations to the prediction of maximum oxygen intake from cardiac frequency measurements. Journal of Applied Physiology 1968; 24: 700-706

DAVIS PO, BIERSNER RJ, BARNARD RJ, SCHAMADAN J. Medical evaluation of fire fighters: how fit are they for duty?. Postgraduate Medicine 1982; 72: 241-148

DAVIS PO, DOTSON CO. Heart rate responses to fire fighting activities. Ambulatory Electrocardiology 1978; 1: 15-18

DAVIS PO, SANTA MARIA DL. Energy costs of wearing fire fighters' protective equipment. International Fire Chief 1975; 41: 10-11

EICHNA LW, BEAN WB, ASHE WF. Comparison of tests of physical fitness. Fort Knox (KY): Armed Ground Forces Medical Research Laboratory, 1944.

ELIOPULOS E, ARMSTRONG BK, SPICKETT JT, HEYWORTH F. Mortality of fire fighters in Western Australia. British Journal of Industrial Medicine 1984; 41: 183-187

ELLIOTT DH, DAVIS J. The causes of underwater acidents. In: Bennett PB, Elliott DH. The physiology and medicine of diving. 3rd ed. London: Bailliere Tindall, 1982: 537-549

EMPLOYMENT MEDICAL ADVISORY SERVICE. The medical examination of divers. London: Health and Safety Executive, 1981. (EMAS Medical Advice Leaflet MA1)

ERKAN N. Hypoxia-exercise stress test (HEST) for the assessment of coronary arterial disease. Ankara: Middle East Technical University, 1977.

FARDY PS, HELLERSTEIN HK. A comparison of continuous and intermittent progressive multistage exercise testing. Medicine and Science in sports 1978; 10: 7-12

FOEHR R. Development of a simplified test to determine individual aptitude for physical effort and exposure to heat. Luxembourg: Commission of the European Communities, 1974. (Community Ergonomics Research Technical Report No.20) (ECSC Doc.No.3825/74)

FRANKLIN BA. Exercise testing, training and arm ergometry. Sports Medicine 1985; 2: 100-119

FROELICHER VF, BRAMMELL H, DAVIS G, NOGUERA I, STEWART A, LANCASTER MC. A comparison of three maximal treadmill exercise protocols. Journal of Applied Physiology 1974; 36: 720-725

GRAVELING RA. Some physiological effects of mental stress and physical activity. PhD Thesis. Salford: University of Salford, 1978.

HEALTH AND SAFETY. Diving Operations at Work Regulations 1981. London: HM Stationery Office, 1981. (SI 1981 No. 399)

HERMANSEN L, SALTIN B. Oxygen uptake during maximal treadmill and bicycle exercise. Journal of Applied Physiology 1969; 26: 31-37

JOHNSON RE, BROUHA L, DARLING RC. A test of physical fitness for strenuous exertion. Revue Canadienne de Biologie 1942; 1: 491-503

KILBOM A. Physical work capacity of firemen. With special reference to demands during fire fighting. Scandinavian Journal of Work, Environment and Health 1980; 6: 48-57

KUORINKA I, KORHONEN O. Firefighters' reaction to alarm, an ECG and heart rate study. Journal of Occupational Medicine 1981; 23: 762-766

LEBLANC C, BOUCHARD C, GODBOUT P, MONDOR JC. Specificity of submaximal working capacity. Journal of Sports Medicine and Physical Fitness 1981; 21: 15-22

LEGGE BJ, BANISTER EW. The Astrand-Ryming nomogram revisited. Journal of Applied Physiology 1986; 61: 1203-1209

LEMON PWR, HERMISTON RT. The human energy cost of fire fighting. Journal of Occupational Medicine 1977; 19: 558-562

LIND AR, MCNICOL GW. The selection of mines-rescue personnel: comments arising from results of a recent accident. Medical Officers' Broadsheet 1968; 8: 1-5

LOUHEVAARA V, TUOMI T, KORHONEN O, JAAKKOLA J. Cardiorespiratory effects of respiratory protective devices during exercise in well-trained men. European Journal of Applied Physiology 1984; 52: 340-345

LOUHEVAARA V, TUOMI T, SMOLANDER J, KORHONEN O, TOSSAVAINEN A, JAAKKOLA J. Cardiorespiratory strain in jobs that require respiratory protection. Int.Archives of Occupational and Environmental Health 1985; 55: 195-206

MANNING JE, GRIGGS TR. Heart rates in fire fighters using light and heavy breathing equipment: similar near-maximal exertion in response to multiple work load conditions. British Journal of Industrial Medicine 1983; 25: 215-218

MASTROMATTEO E. Mortality in city firemen. AMA Archives of Industrial Health 1959; 20: 227-233

MUSK AW, MONSON RR, PETERS JM, PETERS RK. Mortality among Boston firefighters, 1915-1975. British Journal of Industrial Medicine 1978; 35: 104-108

MYLES WS, BROWN TE, POPE JI. A reassessment of a running test as a measure of cardiorespiratory fitness. Ergonomics 1980; 23: 543-547

MYLES WS, TOFT RJ. A cycle ergometer test of maximal aerobic power. European Journal of Applied Physiology 1982; 49: 121-129

NAGLE FJ, BALKE B, NAUGHTON JP. Gradational step tests for assessing work capacity. Journal of Applied Physiology 1965; 20: 745-748

RAVEN PB, DAVIS TO, SHAFER CL, LINNEBUR AC. Maximal stress test performance while wearing a self-contained breathing pparatus. Journal of Occupational Medicine 1977; 19: 802-806

ROBERTSHAW SA, REED JW, MORTIMORE IL, COTES JE, AFACAN AS, GROGAN JB. Submaximal alternatives to the Harvard pack index as guides to maximal oxygen uptake (physical fitness). Ergonomics 1984; 27: 177-185

SCHULTE J. An effective minimum exercise programme for mine rescue teams and similar groups. Results of a pilot study carried out in cooperation with the Cologne College of Physical Education. Medical Service Broadsheet 1983; 22(3-4): 1-22

SHEPHARD RJ. The relative merits of the step test, bicycle ergometer, and treadmill in the assessment of cardio-respiratory fitness. International Zeitschrift fur angewandte Physiologie 1966; 23: 219-230

SPENCER TD. Medical aspects of mines rescue. In: Rogan J, ed. Medicine in the mining industries. London: Heinemann Medical Books, 1972: 285-298

THOMPSON J, BARR D, MCDONALD DR, RENNIE MJ. North Sea divers are no fitter than sedentary men. Letter. Lancet 1984; ii: 107-108

VAN DER MERWE GW, VAN DER HEEVER JAJ, VAN DER WALT WH. A submaximal cardiorespiratory test for the assessment of work capacity. Journal for Research in Sport, PE and Recreation 1980; 3: 36-47

HEAD OFFICE:

Research Avenue North,
Riccarton,
Edinburgh, EH14 4AP,
United Kingdom
Telephone: +44 (0)870 850 5131
Facsimile: +44 (0)870 850 5132

Tapton Park Innovation Centre,
Brimington Road, Tapton,
Chesterfield, Derbyshire, S41 0TZ,
United Kingdom
Telephone: +44 (0)1246 557866
Facsimile: +44 (0)1246 551212

Research House Business Centre,
Fraser Road,
Perivale, Middlesex, UB6 7AQ,
United Kingdom
Telephone: +44 (0)208 537 3491/2
Facsimile: +44 (0)208 537 3493

Brookside Business Park,
Cold Meece,
Stone, Staffs, ST15 0RZ,
United Kingdom
Telephone: +44 (0)1785 764810
Facsimile: +44 (0)1785 764811

Email: iom@iom-world.org