CRITICAL REVIEW

METHODS OF MEASUREMENT OF MUSCLE AND JOINT FUNCTION

NANCY SALTER, OXFORD, ENGLAND

From the Medical Research Council Unit for Research on Climate and Working Efficiency, Department of Anatomy, University of Oxford

"In the movements of an extremity, the two most important factors involved are the power or amount of force present and the amplitude or range of motion; the former being indicative of the amount of function present in the muscles crossing the joint and the latter of the amount of motion between the joint surfaces" (Robinson 1921). It is perhaps necessary to add a third equally important factor, the coordination of muscular contraction required to bring about the desired movement, for without this the limbs are unable to serve a useful function. This review is specifically concerned with the two criteria of joint function mentioned by Robinson, although accurate measurements of muscle strength and amplitude of movement with the voluntary participation of the subject demand normal coordination of the muscle activity.

THE NEED FOR ACCURATE METHODS

Study in this field of medicine received its main impetus from the large number of orthopaedic cases resulting from the 1914–18 war (Nutter 1919, Albee and Gilliland 1920, Clark 1920, Rosén 1922, Silver 1923) and from the poliomyelitis epidemics in the United States at the beginning of the century (Lovett 1916). Investigations carried out at this time indicated the need for accurate assessment of muscle and joint function and led to the development of numerous instruments and methods for measuring and recording the required values. Work was carried on spasmodically until the second world war, when interest in such measurements was renewed because of the increased incidence of limb injuries. Also, with the prevalence of poliomyelitis during the last ten years, further studies have been undertaken in connection with the problems of rehabilitation after this disease.

The practical application of the accurate knowledge of the force and amplitude of movements is apparent in many branches of medicine (Martin and Lovett 1915, Albee and Gilliland 1920, Rosén 1922, Wiechec and Krusen 1939, Molander and Weinmann 1942, Schwab, Watkins and Brazier 1943, Taylor and Brozek 1944, McIntosh, Badgley, Ghormley, Gudakunst, Ivy, Karsner, Lee and Viets 1945, Duvall 1948, Hellebrandt, Skowlund and Kelso 1948, and McBride 1950). In clinical work, for example, accurate measurements are important in diagnosis and prognosis and as a guide to treatment and its effects. They also serve to record the ultimate functional recovery or the residual disability and are particularly useful in deciding when treatment should cease. Repeated measurements during treatment were recommended by Albee and Gilliland (1920) in order to arouse and maintain the interest of the patient in the progress of his own case; under the term "metrotherapy" they included it as a definite part of treatment. Wiechec and Krusen (1939), considering the amplitude of movements, also supported this view. We have found that patients show great interest in frequent accurate measurements of the strength of a muscle group or the amplitude of a movement.

In industrial accident cases, accurate records may be required of the extent of the injury and the degree of functional disability. The just award of compensation may depend on the accurate assessment of the degree of disability (Rosén 1922).

Accurate measurements on normal subjects are of fundamental interest in the study of joint mechanics (Darcus 1951, Salter and Darcus 1952, Darcus and Salter 1953, Salter and Darcus 1953), and may be concerned in specific problems which have their application in clinical work (Darcus 1953). Accurate data are also of value outside the clinical field; normal
values for the strength and amplitude of movements have become of increasing importance in the design of equipment and machinery. Certain tests of muscle strength have also been incorporated in series of "physical fitness" tests used by the Services. Yet for whatever purpose measurements are made, they can only be as useful as they are reliable.

ADVANTAGES AND DISADVANTAGES OF SUBJECTIVE AND OBJECTIVE METHODS

The earliest methods of measuring muscle strength and the amplitude of joint movements were subjective—that is, they were dependent on the personal impressions of an observer. The need for reliable and more accurate methods has led, in many cases, to the replacement of subjective estimation by more objective methods, which make use of an independent means of assessment, thus eliminating so far as possible the human element. However, subjective methods are still used in clinical work, and it has been stated that no amount of objective measurement of work capacity can be substituted for the subjective estimation of an experienced surgeon (Milch 1945). The advantages of objective methods mentioned below, however, warrant the use of quantitative measurement in conjunction with personal impressions wherever possible.

The main disadvantage of a subjective method is that it is crude; there can be no fine discrimination between different levels of muscle activity. For this reason it is of little or no use as a quantitative measure in fundamental or applied problems and has certain disadvantages in clinical work. Impressions gained from observing muscle and joint action may vary from time to time in the same observer and almost certainly do vary in different observers. This may lead to incorrect conclusions, particularly if the rate of recovery in a patient is slow; and it is therefore essential in the personal assessment of joint and muscle function for consecutive tests on any one patient to be made by the same observer. Discussing these disadvantages, Lovett and Martin (1916b) stated: "Impressions that electricity of one kind or another, or rest, or exercise were beneficial have filled the literature; unsupported assertions, marvellous cures and fantastic treatments have too often been advanced on the slenderest of grounds". A further disadvantage of the subjective method is that, since there are necessarily few levels of function that can be distinguished, recorded improvement is slow and the patient therefore derives little encouragement from the tests. An advantage of considering only a few functional levels is that any one grade is wide enough to cover incidental variations.

The main advantages of a subjective method are that it is quick and requires no apparatus, and it can be used for any muscle or joint in individuals of any age. In the assessment of muscle strength, the empirical index of function that it provides has the appearance of being complete in itself and of requiring no comparison with a table of normal values as do numerical results.

In contrast, the objective method may be time-consuming and relatively expensive and yields results which require comparison with a standard. A further disadvantage that may arise in objective testing if an instrument is used for any length of time is that training or fatigue may affect the results. The importance of these factors may possibly be assessed if measurements are taken on the normal side for comparison.

The main advantage of the objective method is that it is reliable and, since measurements can be taken to the required degree of accuracy, the method has a wide application in studies on normal individuals as well as in clinical work. Furthermore, small changes in muscle strength or amplitude of movement can be demonstrated, which is important in providing encouragement to the patient as well as useful information for the clinician. Milch (1945) pointed out that objective methods can provide readings which may be accurately compared with further readings taken on the same individual at different times. This is one of the most important advantages of such methods, since the rate of recovery, effect of treatment, etc., can be assessed.
From the following discussion it will be seen that the general trend is from simple subjective methods to more accurate and complicated objective methods, and that, whereas the former provide valuable information in routine clinical work, more reliable means have been evolved for the accurate study of clinical and applied problems. One point of great importance must be stressed: whether the tests made are subjective or objective, it is essential that the results are recorded accurately and intelligibly.

METHODS OF MEASURING MUSCLE STRENGTH

Probably no single method of determining the strength of individual muscles or muscle groups can be used for every circumstance in which such observations are required. Any method, however, must allow reliable recordings to be quickly and easily made and must be adaptable to the measurement of as many muscle groups as possible.

SUBJECTIVE ASSESSMENT OF MUSCLE FUNCTION

Clinical assessment of muscle strength was originally based on purely subjective methods. The degree of muscle contraction was estimated manually and the muscles graded as normal, partly paralysed or totally paralysed. The partly paralysed group included muscles ranging from those that were nearly normal to those producing only a flicker of contraction. The limitations of this method are obvious. This test was amplified by the estimation of the amount of pressure that had to be applied to prevent a movement occurring. A certain degree of objectivity in grading intermediate levels of activity was provided by adopting the following system for muscles whose action is affected by gravity: None = no evidence of contractility; trace = the muscle can be felt to tighten but cannot produce movement; poor = movement with gravity eliminated; fair = movement against gravity; good = movement against gravity and some resistance; normal = movement against gravity and a greater resistance.

This scheme was first put forward by Lovett (1916), and since that time the different levels of function have been given various numerical values. Legg (1936) and Nelson (1947) used values from 0 (normal, or no paralysis) to 5 (complete paralysis), and Kendall and Kendall (1949) modified the scheme slightly and graded muscle strength as a percentage of normal. The grading most commonly used in Great Britain is that proposed by the Medical Research Council in 1942 (M.R.C. Memorandum No. 7)—namely, from 0 (no contraction) to 5 (normal). This method is widely used in clinical work and has the advantage that it provides an empirical measure of function and can be applied to any muscle group in the action of which gravity is important, and in individuals of any age. However, the method is not clearly defined; for example, a muscle may be able to overcome gravity plus some external resistance through part of its range, but may not be able to complete the range of movement even against gravity alone. Also, in practice, muscles are graded as 5 (normal) when they can function adequately, although in fact they might compare unfavourably with unaffected muscles. Thirdly, no standard techniques are adhered to. The following example is given of two methods of grading the quadriceps muscle which were used in the same hospital: 1) The knee is fully extended and the patient is required to maintain this position against applied resistance attempting to flex the knee. 2) The knee is flexed to a right angle and the patient attempts to extend it against resistance applied to oppose the movement. Finally, the manual testing method is of no use for grading muscles whose action is either not affected by gravity, such as the abdominal muscles, or which do not normally act alone, such as the brachialis. Subjective methods in general are of little or no use for studying the strength of normal muscles under different conditions.

OBJECTIVE METHODS

The methods and instruments to be described have been devised for the measurement of strength of individual muscles or muscle groups and for the assessment of muscular efficiency as a whole. Complicated ergometers for determination of energy expenditure
during prolonged periods of work involving large muscle groups, and instruments designed primarily for specific experimental work are not described; neither is the use of electromyography. The different methods to be described can be roughly classified into four types: spring-balance methods, pressure systems, ergographic techniques and strain-gauge methods.

**Spring-balance methods**—From the subjective methods for grading the functional activity of partly paralysed muscles and muscle groups evolved the first objective method for the large-scale determination of muscle strength. The subjective estimation of the amount of resistance necessary to prevent a movement was replaced by a quantitative measure recorded on a spring balance. One of the main differences between this and other tests of muscle strength that have been proposed is that the part played by the subject is to resist a pull rather than to exert his strength in active effort. This method was suggested and developed by Lovett and Martin (1916a) and was the first simple technique for objective muscle testing to be widely used in clinical work, although the principle of the method had been put forward and a suitable instrument devised by Duchenne in 1863 (Bullard 1886). The method, employing a simple spring balance, was able to show in muscles a slight degree of function not detectable by other means, and more accurate quantitative values could be obtained to demonstrate the course of recovery under different kinds of therapy. A sling adjusted round the limb was attached to an accurate spring balance which measured the resistance applied in order to prevent the limb moving from a standard position. The authors drew up a standard technique for the major muscle groups and compiled a table of "normal" values for varying ages (Martin 1921).

A similar spring-balance technique was used in an experimental study by Haxton (1944) and in clinical work by Lewey, Kuhn and Juditski (1947). The latter workers validated the use of their method by comparing the strength of muscle groups expressed as a percentage of normal, with the amplitude of the electromyogram obtained from maximal electrical stimulation of the nerve also expressed as a percentage of normal. Schmier (1939) used a modification of the method of Lovett and Martin incorporating an ink-writing pointer and appliances for the stabilisation of the patient. Later he further modified and elaborated his method, using a gravity lever scale similar to that used by Mayer and Greenberg (1942) (Schmier 1945). These workers had criticised the technique of Lovett and Martin, stating that a spring balance permits too much subjective variation in the hand of different examiners and too great a range of body motion, with a consequent shift of the angle of force and a variation in the strength of contracting muscles.

Clarke, Elkins, Martin and Wakim (1950), in a study of normal muscles, measured the amount of tension applied to a cable appropriately placed for specific movements as in Lovett's method, with a specially adapted and calibrated tensiometer. This instrument works on the spring principle but is geared to be sensitive to contractions which are almost isometric. Using springs, Rudd (1951) developed a dynamometer which could be used for the measurement of hand and finger strength and also for resistance exercises.

Two instruments commonly used in physical fitness tests as well as in clinical and anthropological work are the hand-grip dynamometer and the oval-spring hand dynamometer. Both these instruments incorporate a spring system but, as well as the disadvantage common to such systems (see below) they are limited in use, for they can be used only to measure the power of the hand grip and they are not adjustable for hands of different sizes. Regnier (1807) (cited by Bullard 1886) was the first to devise a hand dynamometer of the oval-spring type. A more compact instrument was developed by Mathieu and by Robert and Collin about 1870 (Bullard 1886) and has been extensively used with only minor modifications. The disadvantage of using the grip dynamometer as an index of general strength was noted by Martin and Rich (1918): the subject uses one of the most complex musculatures of the body and one which receives much special training. The test therefore reflects conditions to an exceptional
degree. However, the instruments are simple and easy to use and adaptable to a wide variety of test circumstances.

The following disadvantages of a spring-balance system were noted by Wakim, Gersten, Elkins and Martin (1950): a spring has a high internal resistance, a low range of sensitivity and a changing sensitivity. The fact that any spring can only work over a limited range of sensitivity is probably the greatest disadvantage. A further disadvantage in any instrument involving the compression or extension of a spring is that there is an increase in resistance towards the end of the range of movement—that is, at the stage where the muscle is least able to overcome it. A disadvantage of this and the following two methods of objective muscle testing is that they cannot be used to test muscles which, while not completely paralysed, are too weak to produce movement.

**Pressure systems** — Another group of instruments devised for the objective measurement of muscle strength act by means of a system in which pressure changes occur. Most of these were designed to measure the strength of hand grip. The earliest dynamometer of this type, involving compression of a rubber bulb which causes coloured water to rise in a graduated tube, was devised by Hamilton in 1875 (Bullard 1886). Fox (1917) described a similar instrument, which Amar called a "dynamographic pear," in which the pressure change could be read directly off a scale or recorded kymographically. Modifications of this simple apparatus were used by Brahme (1936) and Schwab (1953), while Hellebrandt, Houtz and Kelso (1950) developed a complex instrument incorporating a pressure system, changes in which activated the lifting of weights.

Tarbell (1950) described a "digit myometer" made for measuring small gains in finger strength after injury. Its main advantage appears to be in its simplicity; the strength of the finger is proportional to the increase in height of the column of mercury in an ordinary sphygmomanometer, pressure being exerted on a lever by the finger and transmitted to the cuff through a metal plate. A similar principle was used by Lewis, Pickering and Rothschild (1931). Newman (1949) designed a myometer which could be used for most muscle groups. The instrument consists of a pressure gauge and measures the resistance offered by a muscle in isometric contraction; the reading on the gauge is proportional to the force necessary to overcome the isometric contraction of the muscle under test. The gauge is set in a small cylinder, from one end of which extends a short shaft and a pressure-transmitting button. A built-in hydraulic pressure converter transmits the linear force exerted on the button to the pressure gauge.

**Ergographic methods** — These involve simple graphic representation of repeated muscular exertions against a resistance provided usually by weights or springs. Instead of giving a single value for maximum muscle strength at any one time, this method provides a record showing the rate of onset of fatigue with a load that may be initially maximal or sub-maximal. In this way the endurance of a muscle or muscle group can be tested and the total work done during a known period of time can be calculated, thus indicating the functional capacity of the muscle or muscle group (Molander and Weinmann 1942). A disadvantage of the method is the length of time required for testing, which makes it impracticable for routine clinical work. However, the instrument can also be used for therapeutic exercise.

The original weight ergograph for recording finger contractions was devised by Mosso (1890) for experimental studies on fatigue. The ergograph record was made with a pen carried by a small block which moved along a horizontal track. On flexion of the finger, a string fastened to it by a leather sling moves the block together with weights attached by a cord and passing over a pulley. This principle has been used more recently by other workers in the clinical field (Sonntag 1917, Simon and Simonnet 1938, Kinard and Coleman 1946) as well as for experimental studies (Martin 1921). One objection to the original ergograph is that as the moving part moves in an arc the angle of pull on the weight is continually changing, and therefore it is difficult to calculate accurately the work done by the muscle. Hellebrandt and her co-workers
(Hellebrandt, Skowlund and Kelso 1948, and Hellebrandt, Kelso, Houtz and Eubank 1950) have developed a complex modification of the apparatus incorporating a wheel to the circumference of which is fastened a bar which is activated by the part under study. Thus the angle at which the force is exerted remains constant throughout the range of movement. In this instrument the total distance moved and the number of contractions are recorded independently of the kymograph tracing.

Simple weight lifting has been used in another way as a means of determining muscle strength. Delorme (1945) advocated the use of repeated exertions with gradually increasing loads to find the "one repetition maximum," that is, the maximum load that can be lifted through the full range of movement once only. A disadvantage of taking this as a strength index appears to be that the patient may be fatigued by the time the maximal load is determined (Zinovieff 1951). A further disadvantage of any weight-lifting technique is that there may be dangers in moving heavy loads. If a subject suddenly becomes unable to lift the load, a rapid movement in the reverse direction may cause muscle damage. There is also the danger of falling weights.

Theoretically, the strength of muscles which are too weak to lift even the weight of the limb against gravity can be gauged by determining the weight necessary to assist the muscle and enable movement to take place. In practice, however, it might be difficult to prevent the muscle from playing a purely passive role.

Franz (1901) objected to the weight-lifting type of ergograph because it does not measure isometric muscle contraction. He replaced the lifting of weights by the extension of a spring, and this method was also used by Hough (1901) and Hall (1902) and more recently by Maison and Broeker (1941) and Russell (1952). Some of the pressure-system instruments already described have been used as ergographs as well as for single tests (Brahme 1936, Hellebrandt, Houtz and Kelso 1950, Schwab 1953).

**Strain gauge methods**—More recent methods for the accurate determination of muscle strength involve the use of electrical strain gauges, the advantages of which are their low inertia and high degree and wide range of sensitivity. In such methods the applied muscle force is allowed to deform metal bars or rings to which the gauges are attached. The deformation causes a change in the electrical resistance of the gauges which can be recorded either on a galvanometer or, after amplification, on an oscillograph or by pens. Ralston, Inman, Strait and Shaffrath (1947) used a strain-gauge dynamometer for the measurement of isometric contraction of human upper limb muscles in amputees. Strain gauges were attached, two on the inner and two on the outer side of a ring set in a cable which was fixed at one end and attached at the other end to the tendon of the muscle under study. Tension on the cable produced changes in resistance of the gauges which were recorded oscillographically. Wakim et al. (1950) used this method on normal muscles, the cable being attached by a sling to the extremity being tested. They found results similar to those obtained using the cable tensiometer (Clarke et al. 1950). Tuttle, Janney and Thompson (1950) used strain gauges in a hand-grip dynamometer, and Darcus (1953) described a strain-gauge apparatus for the measurement of rotary isometric torques about various joints.

The latter apparatus has recently been accepted for use in a Medical Research Council investigation on the effects of different exercise routines on the strength of weakened muscle groups following poliomyelitis. This dynamometer appears to be highly suitable both for research projects and for routine clinical testing. In clinical work, however, it would supplement rather than replace the 0–5 muscle charting system.

**MEASUREMENT OF THE AMPLITUDE OF JOINT MOVEMENT**

The requirements of an "ideal" method for determining the amplitude of joint movements are that it should allow the recording of accurate readings which are reproducible by different observers, and that it should be applicable to all limb joints in people of different
body dimensions. If the method demands an apparatus, this should be simple in construction and use and should not interfere with joint movement. Most of the methods so far devised fulfil only a few of these requirements.

The variety of methods described may be broadly divided into subjective assessments and objective measurements with various types of apparatus.

**SUBJECTIVE ASSESSMENT**

Subjective assessment, which is dependent on visual approximation, was the first method used and is still commonly employed in clinical assessment of joint function. An adaptation of the visual method was provided by Cleveland (1918). He used circular charts on to which the limits of movement of the moving limb itself, estimated by visual means, were transferred. In the centre of the charts were diagrams of the part concerned showing the axis of movement.

Clark (1921) used a more refined modification of the visual method for the measurement of rotary movements. He used a protractor scale arched over the distal end of the rotating limb, with the limb itself serving as an indicator.

**OBJECTIVE METHODS**

To allow measurements to be made accurately and to be independent of the observer, numerous instruments have been devised during recent years. These have been termed "arthrometers" (instruments for measuring joints), "goniometers" (instruments for measuring angles) or "fleximeters" (instruments for measuring the degree of bending). These may be classified roughly into three types: protractor and pendulum arthrometers and independent methods.

**The protractor arthrometer**—This consists essentially of two rigid shafts intersecting at a union allowing movement at right angles to their longitudinal axes. A protractor is fixed to one of the shafts so that its centre corresponds with this union; the other can move independently of the protractor and acts as an indicator. The shaft to which the protractor is attached is held or strapped parallel to the longitudinal axis of one part of the limb, the union of the shafts corresponding to the axis of the movement being studied, and the other is held against or strapped to the other part of the limb. The degree of movement can then be measured directly from the protractor scale. Many modifications and improvement of this basic pattern are reported in the literature (Camus and Faidherbe 1915, Manouvrier 1915, Alquier 1916, Fox 1917, Dauzet 1919, Albee and Gilliland 1920, Marble 1920, Rosén 1922, Silver 1923, Conwell 1925, Cobe 1928, Parker 1929, West 1945, Dorinson and Wagner 1948). All have been used to measure flexion and extension movements, particularly at the knee and elbow.

Although the error inherent in the construction of these protractor arthrometers is negligible, inaccuracies occur during their use, mainly from their faulty application. Such faults include failure to ensure that the shafts of the instrument are parallel to the longitudinal axes of the corresponding limb segments, and that the axis of the instrument coincides with the axis of the joint. Wiechec and Krusen (1939), however, considered that with a moderate amount of care in positioning, a simple form of protractor arthrometer, such as the one described by Clark (1920), could attain as great a degree of accuracy as a much more complicated instrument.

Wilson and Stasch (1945) have modified the technique of the measurement of joint movement by the protractor method. They made double-exposure photographs of the arthrometer and the limb, first with the joint in the position of maximal extension and then with the joint in the position of maximal flexion. The amplitude of movement was read off from the photograph, which also provided a permanent record. They also used this photographic method without the arthrometer, determining the amplitude of movement on the developed print by running two straight lines from the axis of the joint along the longitudinal axis of
the movable part of the limb in the two positions, and measuring the angle with a protractor. This latter method was found to be particularly useful in observations on the wrist, metacarpophalangeal and interphalangeal joints. These methods do not seem to have any particular advantage over other protractor methods, apart from providing a permanent graphic record, and this is outweighed by the increased complexity of the procedure.

**The pendulum arthrometer**—This type of instrument consists of a circular scale, to the centre of which is attached a pointer weighted at one end so that it remains vertical while the scale rotates around it. The scale of the instrument is attached to the movable limb, and during movement of the joint the scale moves around the stationary pointer, thus registering the amplitude of movement. Falconer (cited by Fox and van Breeman 1934) appears to have been the first to devise such an instrument for measuring the amplitude of movement of the knee and elbow joints. Similar types for general use have been described by Glanville and Kreezer (1937) and by Hand (1938).

**Independent methods**—A number of methods which are to a certain extent independent of the subject on whom measurements are taken have been used to study the amplitude of movements. With such methods there is no possibility of the apparatus interfering with normal movement. Two methods are suitable for clinical use.

Wilmer and Elkins (1947), dissatisfied with other methods of measurement of joint movement, devised an arthometer embodying optical principles. This instrument consists of a transparent circular scale attached to a concave reducing lens. On looking at the limb through the lens and scale, the reduced image of the limb appears on the scale. The skin is marked to indicate the axis of movement of the joint and define the longitudinal axes of the bones forming the joint. The limits of movement are marked on the scale by movable pointers. This method is claimed to have the advantages of simplicity, lack of interference with the movement, and adaptability.

Recently Zankel (1951) has described another method which he terms "photogoniometry." A protractor scale is projected on to the limb of the subject so that the centre of the scale is focused on the axis of the joint. The amplitude of movement is measured from the projected radii of the protractor scale, which coincide with the longitudinal axes of the bones articulating at the joint under consideration. Points may be marked on the skin to facilitate the readings. He claims that it is an improvement on previous methods because it is more accurate and more susceptible to duplication by different observers, although he gives no evidence for this. He also contends that this method lends itself better to teaching because many different observers can study the same joint measurements concurrently.

Other methods that have been put forward involve radiographic and cinematographic analyses of movement, but these were not intended for routine clinical assessment of amplitude.

Bakke (1931) studied movements of the vertebral column by serial radiographs, and Harris and Joseph (1949) used radiography to measure the degree of extension of the metacarpophalangeal and interphalangeal joints of the thumb. Weddell and Darcus (1947) and Darcus (1948) studied the amplitude of composite movements of the neck and spine by the analysis of a ciné-film record. Cinéradiography has been used to study movements of the lower limb (Barnett and Napier 1952) and may be of value in determining the amplitude of complex movements.

A review of methods used to measure the amplitude of rotatory movements of the upper limb was given by Darcus and Salter (1953).

In conclusion it can be seen that, despite the efforts made by numerous workers to devise an arthrometer fulfilling all or nearly all of the requirements, there is still no widely accepted method. If used with sufficient care, it would seem that the simple protractor type of arthrometer is the most generally useful, since it provides adequate information on most joints. Where more precise measurements are necessary, a radiographic method is to be recommended.
FACTORS TO BE OBSERVED IN THE USE OF APPARATUS AND METHODS DESCRIBED

In connection with the use of both objective and subjective tests for the measurement of muscle strength and the amplitude of free movements, there are certain factors which merit attention and consideration. The following points are put forward to illustrate the confusion that is liable to arise when, for example, different terminologies are used, and to provoke suggestions and perhaps some measure of standardisation which would be of benefit in this field of study. It is hoped also that the discussion will demonstrate the importance of indicating in case records and published results the exact technique and procedure of a test so that adequate data are available for those wishing to repeat the study or to use the results for comparative purposes. The factors considered here, although there may be others equally important, are posture, test procedure, standards for comparison, nomenclature and normal variability.

Posture—In the assessment of both muscle strength and the amplitude of joint movements, it is necessary to standardise, or at least to specify, the posture of the body and the precise position of the joints under study as well as those of associated joints. It is essential that these positions should be the same on successive occasions on which measurements are made in order to be able to compare readings taken from the same part in the same individual at different times. The position of the body—for example, sitting or lying down—is important from the point of view of comfort and relaxation of the patient and in so far as it affects the position of the limbs. Fox (1917), stressing this point in the measurement of the amplitude of movement, gave as an example the increased flexion of the hip allowed by flexion of the knee which reduces the "ligamentous" action of the hamstring muscles. Similarly the amplitude of dorsiflexion of the foot is restricted by extension of the knee due to the "ligamentous" action of the gastrocnemius and the amplitude of flexion of the metacarpophalangeal joints of the hand by flexion of the wrist due to the ligamentous action of the long extensors. The results of Cobe (1928) show also that the position of the radio-ulnar joints influences the amplitude of abduction and adduction of the wrist.

Experimental studies on the effect of joint position on muscle force (Franke 1920, Hansen and Lindhard 1923, Hvoslev 1928, Garry 1930, Muller 1935, Haxton 1945, Dern, Levene and Blair 1947, Hugh-Jones 1947, Clarke, Elkins, Martin and Wakim 1950, Wakim, Gersten, Elkins and Martin 1950, Wilkie 1950, Darcus 1951, Salter and Darcus 1952) show how important is the consideration of individual joint position in the assessment of muscle strength, whether this assessment is subjective or objective, and for clinical or fundamental work. Posture is also important in the use of the simple and commonly used spring-balance dynamometers, and results of different workers may be of little value for comparison unless standard positions are adopted. This is important, for example, in the "lumbar pull" test which is widely used to obtain an index of general strength. In this test the subject holds the dynamometer handle and pulls vertically upwards, keeping his arms and legs straight. Vernon (1924) had shown that there is an optimum height for the handle below or above which the force diminishes considerably, and after experimental investigation Bedford and Warner (1937) found that the greatest pull was exerted when the handle was at the height of the subject's finger tip.

No standard posture has been advocated for the use of the hand-grip dynamometer. The subject is usually instructed to hold the instrument where he feels that he can exert his greatest force. Fisher and Birren (1946) found that supporting the instrument with the other hand affected the results considerably. It is possible that using a standard posture for the test may also affect the results.

In all muscle testing, both for strength and amplitude of movements, it is necessary to specify the part played by the force of gravity in assisting or resisting movement. It is also necessary to prevent, or at least to be able to detect, substitution or "trick" movements so that one is sure what muscle or muscle groups are acting.
Most accurate results of objective muscle testing would be obtained in the positions of associated joints which normally allow the greatest force to be exerted. This information is available for certain muscle groups (Clarke et al. 1950, Darcus 1951).

**Test procedure**—In all muscle tests the question arises of how many observations should be made on a single occasion. More than one reading is usually taken to avoid chance inaccuracy. Hellebrandt, Parrish and Houtz (1947) showed that an increase in the strength of muscles, which had undoubtedly occurred, was not reflected in tests involving only one reading on a spring dynamometer. They attributed this to the many uncontrollable variables which may influence an isolated observation. However, a study of the reliability of the single muscle test was made by Duvall, Houtz and Hellebrandt (1947) and they found that a single reading is as reliable as the best of three or the best of ten. For clinical measurements of joint amplitude three readings are usually taken. Common methods in the measurement of muscle strength are to take the mean of two or three observations or the best of three; these probably evolved as being the most convenient and the least time-consuming. Doubting the reliability of the "best of three" method, Fisher and Birren (1946) tested the following procedure with a hand-grip dynamometer: scores were taken at three-second intervals with the subject starting at 27 kilograms and increasing by increments of 3 kilograms until his limit was reached. The load increment and the time interval were chosen arbitrarily. Re-test performance showed this to be a reliable procedure and the authors gave it the following advantages: it allows a "warming up" to take place, and the cumulative fatigue ensures the final effort being made under a standard stress. A similar technique was used by Delorme (1945) in clinical work; his subjects worked with a gradually increasing load until a standard number of contractions could not be attained. In considering the number of readings to be taken, the relative importance of fatigue and "warming up" must be noted. The best time interval between successive recordings must be determined, and we have found that five maximal voluntary exertions spaced at intervals of one minute produce little or no fatigue. The speed of making a single exertion—that is, the time over which the maximal force is built up in the muscle, must also be considered. Studies carried out by Hill (1922) indicate that the work done in a single maximal effort increases with increasing time taken to build up the contraction—that is, with decreasing speed of contraction. This is due to the fact that the more rapidly a muscle shortens the more potential energy is used up in overcoming the viscous resistance of the muscle. However, there is an optimal economical speed of working related to maximal efficiency.

In ergographic techniques the question arises of the rate at which maximal contractions should be made. This is determined partly by the movement under consideration and partly by whether or not the muscles are normal. Common time intervals are one, two or three seconds. One bout to complete exhaustion may be recorded or a standard number of contractions may be performed and repeated after intervals of known duration. Foltz, Ivy and Barborka (1942) recommended the use of double work periods, each to exhaustion, with a standard rest pause between. The advantage of this method is that, as well as the total work done, the percentage recovery is obtained.

**Standards for comparison**—In clinical work some standard for comparison is required when measuring either the force in a paretic muscle or the limited amplitude of movement of a joint. As a standard the corresponding measurement taken on the contralateral limb may be used, if this is normal, or reference may be made to a relevant table of "normal" values. So-called normal values for the amplitude of various movements are to be found in the literature (Clark 1929, Gilliland 1921, Rosén 1922, Silver 1923, Frescoln 1929, Sinelnikoff and Grigorowitsch 1931, Coulter and Molander 1935, Glanville and Kreezer 1937, Dorinson and Wagner 1948). Gilliland's values were obtained from "100 observations on male students." However he does not specify the number of observations on each joint. Glanville and Kreezer (1937) studied a number of joints in ten subjects, and the values given by Sinelnikoff
and Grigorowitsch were obtained from 396 male and 118 female subjects. The other workers mentioned give no indication of how their values were obtained or of the size and characteristics of the group of subjects. Such "average" or "normal" values are of limited use, particularly if the range of values which can be considered normal is not given, and, in fact, the values given by different authors for the same movement differ from each other considerably (Darcus and Salter 1953, Salter and Darcus 1953). A more accurate yardstick for the normal amplitude of movement of a joint is probably the contralateral limb, when this is available for study. That there is little difference between the normal amplitudes of certain movements on the right and left sides has been stated by Gilliland (1921). Cobe (1928) and Hewitt (1928), studying the amplitude of movements at the wrist, found differences between right and left sides of over 200 per cent in some cases, but bigger differences were found between individuals, and variations of over 100 per cent were recorded in taking the same measurement on the same individual on different days, suggesting faulty apparatus or technique. Patrick (1946) stated that the range of pronation and supination in both forearms of a single individual is identical. Darcus and Salter (1953) found that the mean difference between the amplitudes on the right and left sides in eight normal subjects was 3 per cent, in contrast with a difference of 20 per cent between subjects.

Considering the measurement of muscle force, Lewey, Kuhn and Juditski (1947) put forward the following objections to using the contralateral side as a standard for comparison: 1) the time of the examination is doubled; 2) there may be a compensatory increase in muscle strength on the normal side; and 3) values on the right side may be higher than on the left side in right-handed subjects. Although the time of testing is doubled, readings taken on the normal side are an advantage in that they reflect normal variations (Darcus 1951, Salter and Darcus 1952). Variations due to psychological, physical and environmental factors can be taken into account in assessing the affected side, as can training and fatigue. At present, comprehensive tables of normal values suitable for comparison under particular test conditions do not exist. Individual variation is so wide and dependent on so many factors (age, sex, physical characteristics, occupation, etc.) that a great deal of work would be necessary to compile adequate information. Using such information, one can only say that a patient falls within the limits of normal variation for his age and sex group or is approximately x per cent of "normal." Lewey et al. (1947) gave a table of normal values for twenty-three muscle groups compiled from observations on a random sample of hospital patients showing no nerve or muscle weakness. They give mean values and standard deviations but omit to give the number of subjects used and their ages, merely stating that the group includes manual and clerical workers and both right- and left-handed persons. They claim to have obtained more satisfactory results using this table than by using the contralateral limb as a yardstick, despite the fact that overall variation between different individuals appears to be much greater than that between right and left sides in a single individual (Provins 1955, Provins and Salter 1955).

Nomenclature—To be of value, records must be uniform and if possible self-explanatory. Difficulties involved in assessing early work on the measurement of joint amplitude included the wide disagreement in nomenclature, which made results liable to misinterpretation. Each author adopted his own nomenclature, often without specifying what he meant by the terms used.

Most workers appear to agree that all measurements should begin from a neutral point, but there is considerable difference of opinion as to where in the whole arc of movement this should be. Usually it is placed at an assumed mid-point or at one or other end of this arc. If the neutral point is at the assumed mid-point, differences are found in the angular value given to this point and in the system of denoting degrees of movement on either side of it. The difficulties of interpretation arising from this lack of consistency are indicated by the figures given by different authors for the amplitude of flexion and extension of the wrist (Table I).
The meaning of these figures is apparent only when their graphic representation is studied (Fig. 1).

Further confusion arises through the different methods of describing a movement. For instance, in joints allowing extension beyond the point at which the bones are in line, the movement beyond this point is sometimes referred to as hyperextension. Abduction and adduction of the wrist are also described as radial and ulnar deviation or flexion. Movements of the shoulder joint are either defined in relation to the main axes of the body or to the axes of the scapula (Johnston 1937).

The first attempt to clarify the nomenclature was made by Clark (1920). He suggested that the angle included between the bones on either side of the joint should be used to express the limits of motion or angle of deformity; thus complete extension of the knee or elbow would be 180 degrees. However, modification of and additions to this scheme for the measurement of abduction, adduction and rotation made it unnecessarily complicated, and it was not generally accepted. Rosén (1922) believed that measurements from a neutral point were not satisfactory and suggested that the angle subtended at the joint axis by the two extremes of excursion of the joint was more suitable. Silver (1923) proposed that "mensuration of the degrees of motion in joints must start from a zero plane. The zero plane should be considered to be an extension of the plane of the long bone or bones immediately proximal to the joint in question." This method had previously been used by Robinson (1921). Cave and Roberts (1936) merely stated that "all movements should be measured by degrees from a neutral point or zero which must be defined." West (1945) measured all movements on a circular scale graduated from 0–360 degrees, using the position of the joints in the anatomical position as a neutral position. Flexion was measured towards 0 degrees and extension towards 360 degrees.

Dorinson and Wagner (1948) put forward another suggestion in order to "facilitate recording and obviate any misinterpretation." They state that "if, for an example, a patient has a 10 degrees' flexion contracture with a range restricted to 40 degrees, this would be obvious by the recording of 'extension-flexion' 170–130 degrees." So far as is known, this suggested method has not been applied to any extent.

It is evident that there is still as much disagreement as ever concerning nomenclature, and that until a universal method is adopted each author must specify exactly which angle he is in fact measuring and which neutral point he is using. To avoid ambiguity, it is stressed that graphic representation of the results should supplement the written description (Cleveland 1918, Nutter 1919, Parker 1929). It is also important to state whether the movements measured are active or passive. The results of Glanville and Kreezer (1937) indicate that the difference between the amplitudes of active and passive movements in normal individuals

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Figures of Different Authors for Amplitude of Flexion and Extension at the Wrist</strong></td>
</tr>
<tr>
<td><strong>Joint</strong></td>
</tr>
<tr>
<td>Clark (1920)</td>
</tr>
<tr>
<td>Robinson (1921)</td>
</tr>
<tr>
<td>Rosén (1922)</td>
</tr>
<tr>
<td>Silver (1923)</td>
</tr>
</tbody>
</table>

* These figures were not intended to represent the complete amplitude.
may be 10 per cent or more, and it is felt that this question deserves further investigations. In clinical conditions the difference may be much larger and, while this is well known, no objective clinical study on the subject has been found.

Similar confusion is less likely to result from work on the measurement of muscle force, providing that the units and the point of application of force are clearly defined, so that only strictly comparable values are in fact compared. Static activity cannot be compared directly with dynamic work because the units are different. Starr (1951), however, has devised a mathematical method of expressing the two components in experimental weight lifting in the same units.

In manual muscle testing in Great Britain muscles are graded numerically from 0 to 5, 0 indicating no evidence of contractility and 5 a normal muscle (Lovett 1916, Medical Research Council Memorandum No. 7, 1942). It was confusing to find in two American papers the same
functional levels given different numerical values (Legg 1936, Nelson 1947). These numerical values may be misleading as they bear no relationship to absolute levels of muscle strength. For example, a muscle of grade 4 is not twice as strong as a grade 2 muscle.

In order to avoid difficulty in the interpretation of results, it is stressed that the exact procedure should in the first instance be described.

**Variability**–Considerable variation is found in measurements of muscle force and of amplitude taken on different occasions in the same individual (Darcus 1961, Salter and Darcus 1953). Such variations will occur in injured as well as normal individuals, and the clinician will wish to know how big a change in muscle force or amplitude must be before it reaches practical significance. A dramatic and maintained improvement offers no problem, and small changes which are insignificant in themselves become significant when they are reflected in the trend of a series of tests made on successive occasions. However, in order to increase the accuracy of assessing any improvement, normal variability should be minimised so far as possible.

Of environmental factors, temperature and possibly humidity and barometric pressure may affect muscle strength. Probably temperature is the most critical, though it can most easily be controlled. Lovett and Martin (1916) noticed that during “the great heat of September 1915” there was a decrease in muscle strength in several persons whose normal values were known. The temperature of limbs is important in testing muscle strength, particularly after poliomyelitis. Lombard (1892) concluded from his experiments that the time of day and barometric pressure influences muscle strength, and the latter has also been suggested by Fischer (1947). These factors have not been proved to have an effect, but the time at which tests are made could perhaps be standardised for each subject.

Distractions to the patient and to the observer should be avoided, concentration on the test maintained and motivation standardised as far as possible. In clinical work the incentive is usually high, particularly if the patient is allowed to watch the readings being taken and to follow the results of the test (Albee and Gilliland 1920). The willingness to exert a maximal effort when this is required may markedly affect the results, and there is no simple objective method of distinguishing between the physiological and psychological end points of effort. The interest and co-operation of the subject should be sought. This is particularly important in experimental work with normal subjects, because boredom and antagonism would give inaccurate results and wide variation within groups of readings. Individuals being tested should be instructed to relax “mentally” as well as physically, for apprehension or anxiety may affect the degree of muscle tension (Jacobson 1946, Barlow 1947, Lundervold 1952). The necessity for rigid control of environmental, physiological and psychological factors is emphasised by Taylor and Brozek (1944). Factors which could be to some extent controlled in hospital patients, but which are difficult if not impossible to control in other subjects, are the amount of rest and the amount and type of activity and exercise. These factors may affect readings taken at long intervals or over extended periods. Pain may limit the degree of muscle contraction in clinical cases and must be taken into account when relevant. If one side of the body is not affected by disease or injury, control tests on this side might help in distinguishing between the effect of environmental factors and factors acting on the affected side alone. Similarly, contralateral tests might assist in the distinction between a training effect in the remaining innervated fibres in a paretic muscle and improvement perhaps due to functional re-innervation of previously denervated fibres.

In summing up this discussion of the importance of test procedure, posture, nomenclature, standards for comparison and normal variability in the measurement of muscle and joint function, it can be seen that a consideration of these factors is essential if the work carried out is to be as useful as possible. Although it is impossible to apply rigid techniques, because modifications are required to meet individual needs, it is considered that there is room for some standardisation, particularly of nomenclature. Until there is a generally accepted system
of denoting the angular position of joints it is necessary for workers to describe accurately the system they use as well as specifying details of posture and test procedure. It is appreciated that time is valuable in routine hospital work, but the little extra time required to specify essential details of the procedure would be well spent. If this is done, future presentation of data will be clearer and more useful than it has been before.

**SUMMARY**

1. The importance of accurate methods of measuring the strength of muscles and the amplitude of joint movements in man, both in clinical fields and as criteria of normal function, is discussed.
2. The advantages and disadvantages of subjective and objective methods are reviewed.
3. The main types of apparatus used for the assessment of muscle strength in both normal and clinical conditions are described. A dynamometer of the strain-gauge type is recommended.
4. Methods of measuring the amplitude of movements in man are also described. The protractor type arthrometer is thought to be the most suitable for routine clinical work, but for research purposes a radiographic method may be preferable.
5. The following factors, which must be considered if the measurements taken are to be of greatest use, are discussed: posture, test procedure, standards for comparison, nomenclature and normal variability.

**REFERENCES**


Fox, R. F. (1917): Demonstration of the Mensuration Apparatus in Use at the Red Cross Clinic for the Physical Treatment of Officers. Proceedings of the Royal Society of Medicine (Section of Balneology and Climatology), 10, 63.


Vol. 37 B, No. 3, August 1955


M.R.C. War Memorandum No. 7 (1942): Aids to the Investigation of Peripheral Nerve Injuries.


METHODS OF MEASUREMENT OF MUSCLE AND JOINT FUNCTION