The Foundation of Clinical Decisions

Decision-making is ... something which concerns all of us, both as the makers of the choice and as sufferers of the consequences.

Lindley

And what does 'outgrabe' mean? Well, 'outgribing' is something between bellowing and whistling, with a kind of sneeze in the middle: However, you'll hear it done, maybe – down in the wood yonder – and when you've once heard it you'll be quite content ...

Lewis Carrol in 'Through the Looking-Glass'

A person who feels ill will usually seek medical advice, and the doctor, having listened to her patient’s complaints, will make those decisions which, to the best of her knowledge, will help the patient most. This sequence of events is not new. If a patient at the beginning of the thirteenth century had developed an acute fever, the physician would have prescribed some medicinal herb. White benedicta (blessed thistle) might have been the choice, because this herb was said to possess great healing powers when it was taken on an empty stomach and when Pater Noster and Ave Maria were recited three times. If the incident had taken place 600 years later the doctor might have made a diagnosis of pneumonia using the newly invented stethoscope, and would probably have ordered blood letting, customary dietary measures and blistering (induced by dried, pulverized Spanish fly).

The situation today is just the same, except that the doctor has a choice between many more investigations and treatments, and that the decision may
have much greater influence on the course of the disease, and, therefore, on the future life of the patient. The decision to treat a patient suffering from pneumonia with penicillin may ensure her survival when otherwise she might have died, and the decision not to do a lumbar puncture in a febrile patient may lead to her death from meningitis, although she could have been saved.

The clinical decision process

It is not so simple, however, that all positive decisions are beneficial and that all negative decisions – omissions – are harmful. All active treatments can produce harm (otherwise, they wouldn’t be active) and many diagnostic procedures (e.g. liver biopsies and endoscopic examinations) are unpleasant and may cause complications. The clinician must carefully consider the consequences of her actions, both for the individual patient, and, as we shall discuss later, for the health service as a whole.

The clinical decision process is complex, but may be illustrated by a simple flow chart (Fig. 1.1). When contact between patient and doctor has been established, the data collection begins (Step 1). The doctor interrogates the patient, does a physical examination and asks for appropriate blood tests, X-ray examinations etc. When the examinations are concluded, the clinician assesses the data that have been collected and tries to make a diagnosis (Step 2). To do this she uses her nosographic knowledge, i.e. her knowledge of the manifestations of different diseases (nosography = disease description, derived from Greek nosos = disease).

A diagnosis may be more or less certain, and the clinician has to ask herself whether or not the diagnosis is sufficiently well founded to proceed to treatment (Step 3). If the answer is ‘no’, the process returns to Step 1 and the investigations continue. If the answer is ‘yes’ the clinician proceeds to Step 4.

At this point she must once again draw on her nosographic knowledge, this time of the prognosis of the disease and the effect of different treatments. She chooses the treatment that is considered likely to help the patient most, and if the patient progresses as expected, the process comes to an end (Step 5).

This presentation is, of course, greatly simplified and frequently the decisions proceed in a different way. Sometimes the patient does not respond to treatment as expected, and the diagnosis must be revised; and sometimes it is necessary to institute treatment before the final diagnosis is made, as, for instance, in cases of haemorrhagic shock when treatment is started before the site of the bleeding is known. The flow chart also ignores the fact that in chronic diseases clinicians
must consider the long-term effects of their treatments, and the presentation does not take into account that doctors concern themselves not only with treatment, but also with prophylaxis.

Therefore, the flow chart in Fig. 1.1 is by no means universally valid, but it may serve as a framework for a systematic analysis of the decision process.

Clinical reasoning may be deductive or empirical. Clinicians reason deductively when they base their treatment decisions on deductions from theoretical knowledge of disease mechanisms and the mechanism of action of different drugs, whereas they reason empirically when their decision is based on experience that has been gained from the treatment of other patients. When a clinician recommends a β₂-agonist for the treatment of asthma, her reasoning is deductive if she argues that the symptoms are caused by bronchoconstriction and that β₂-agonists decrease this. Her reasoning is empirical, however, if she recommends steroid inhalations because randomized trials have shown a good and sustained effect with little harm.

Deductive and empirical reasoning constitute the scientific component of clinical decision-making (if we use the word scientific in its narrow sense, i.e. 'pertaining to the natural sciences'), but added to this is the humanistic
component which comprises reasoning based on an *understanding* of the patient as a fellow human being, and *ethical* reasoning based on ethical norms (see Chapter 7). Thus, clinical decision-making is a synthesis of four types of reasoning.

**Clinical data**

A house physician at a medical unit reports to a registrar that she has just seen a new patient who presents with a red and swollen left lower leg and tenderness of the calf. The registrar accepts the suggestion that anticoagulant treatment is instituted in order to prevent progression of the surmised deep venous thrombosis. However, another houseman who has also seen the patient objects. It is true that the leg is red and swollen, but the demarcation of the erythema is sharp, the affected skin is raised compared with the adjoining normal skin, and the patient has a small ulceration on the foot. The registrar now correctly diagnoses erysipelas and changes the treatment to penicillin. The example illustrates the well-known danger involved when decisions are made by a doctor who has not examined the patient herself. Both diagnosis and treatment depend on the collected data and no amount of professional knowledge can compensate for incorrect information. Therefore, it is not possible to attempt an analysis of diagnostic and therapeutic decision-making until the information that is used for the decisions has been analysed in detail, but unfortunately this analysis is hampered by the fact that many of the terms that we generally use are vague and ill defined.

In this book I shall use the term *clinical data* to denote all those data about the individual patient which are relevant for the decision process, and collectively these data are said to constitute the patient’s *clinical picture*. Consequently, the clinical data in this wide sense comprise both the data recorded at the bedside (symptoms and signs), i.e. the truly clinical data, and the results of laboratory investigations, i.e. the so-called paraclinical data. Further, the clinical data (and the clinical picture) include negative findings, such as ‘the lack of neck stiffness’ in a febrile patient.

I shall use the following definitions of the different types of clinical data:

- *Subjective symptoms*. These are the sensations noted by the patient (e.g. pain, clouded vision and dizziness) and the patient’s mood (e.g. depression and anxiety).
CLINICAL DATA

- **Objective symptoms.** This term signifies all observations made by the patient or the relatives concerning the patient’s body and its products, e.g. swollen ankles, blood in the urine or an epileptic attack.

- **Physical signs.** These comprise all those observations that are made by the doctor during the physical examination, e.g. a cardiac murmur, swollen lymph nodes or jaundice. Some of the recorded ‘signs’, such as tenderness, dysaesthesia or loss of central vision of one eye, fall into a special group. They are subjective symptoms that are only noticed by the patient during the physical examination, and they may appropriately be called provoked symptoms.

- **Paraclinical data.** They include all laboratory results, and the results of all examinations not done by the clinician herself, such as blood analyses or radiological and histological findings. Paraclinical data may be either descriptive, e.g. the shadow on a chest X-ray, or quantitative, e.g. the blood glucose concentration.

The patient’s record will also contain other data that may be of paramount importance, e.g. information about occupation, family life, previous illnesses, medication, smoking, drinking and other habits.

The erysipelas case illustrated that incomplete or unreliable information may easily lead the decision process astray, and it is worthwhile considering how the clinical data come to the clinician’s attention. We may for this purpose distinguish between three types of data:

1. The symptoms that make the patients seek medical advice;
2. The data that are revealed by the routine questioning and the routine investigation of the patient;
3. The data that are the results of diagnostic tests that are carried out to confirm or exclude various diagnostic possibilities.

The first type of data may be labelled the iatrotropic symptoms (from Greek iatros = doctor and trope = turn). They are the subjective and objective symptoms that make the patient turn to her doctor as opposed to the non-iatrotropic symptoms which are only disclosed during the taking of the history. In the same way one may distinguish between iatrotropic and non-iatrotropic cases of a disease. A patient who sends for a doctor because of a high temperature and who, during history taking admits having epigastric pain, may represent an iatrotropic case of pneumonia and a non-iatrotropic case of duodenal ulcer.
Non-iatrotropic cases may also be diagnosed during a routine medical ‘check-up’ or at mass screening for some disease, e.g. tuberculosis.

Iatrotropic symptoms are of particular importance as they usually represent that problem which the doctor, in the eyes of the patient, has to solve, and they ought to be given particular prominence, especially in hospital notes. Nowadays it is not rare in hospital practice that the investigations bring some unexpected findings to light which then lead to further investigations along a side-track. After a while the whole staff is interested in, say, the immunoglobulin pattern, and nobody remembers why the patient was admitted. Only on the day on which the patient is discharged will she say, ‘But you have not done anything about my backache!’ The advanced specialization of hospital departments invites the occurrence of such incidents. If a patient presents a complex clinical picture, the subspecialized physician may more or less consciously emphasize those clinical data which pertain to her own field of interest, whereas the remaining data are treated more lightly.

A symptom may become iatrotropic for a number of reasons. One patient with abdominal pain may ask to see her doctor because she is afraid of cancer while somebody else with the same complaint may be worried about losing her job, and in other cases the reason for the visit is not directly related to the symptoms. The patient may have felt a lump in the breast which she dares not mention, but hopes that the doctor will find, or she may have problems at work or at home which made her contact her doctor on the pretext of some mild symptom, which under normal circumstances she would have accepted. Personal problems of any kind may lower the threshold of iatrotropy.

The second type of data are recorded routinely from all patients. In hospital practice they comprise answers to standard questions during history taking, the results of the ordinary physical examination and some simple paraclinical tests, such as haemoglobin determination and urine analysis. The routine history taking and examination are to a large extent determined by tradition and from time to time they must be brought up to date. It is no longer necessary in some countries to ask all elderly patients whether they have had rheumatic fever or diphtheria, but it is important to ask detailed questions about their social network and their living conditions. Perhaps the dangers of working with organic solvents would have been detected earlier if the notes contained more routine information about occupation and working conditions.

The third type of data are those collected during the diagnostic process, which begins as soon as the iatrotropic symptoms have been recorded. The clinician will, for instance, ask a jaundiced patient if she has had abdominal pain and or if she has travelled abroad, and such specific questions may well
be likened to diagnostic tests, which aim at confirming or excluding different diagnostic possibilities. In other words, the process passes through the loop in the flow chart (Fig. 1.1) many times already during the taking of the history.

Scales of measurement

Clinical data have many characteristics that are not peculiar to medicine, and I shall now consider the classification of data in general terms, using nonmedical examples. There are three levels of measurement scales: the nominal scale, the ordinal scale and the interval scale.

At a music lesson, a recording of a short passage from an orchestral work is played and the children are asked to identify the solo instrument at a particular moment. Each child is asked to indicate her reply on a list of all the instruments of the orchestra. Such a list, which is used to classify qualitative observations in a series of named categories, is called a nominal scale, which, from a formal point of view, must fulfill three conditions. Firstly, each category or class must be well defined, and we shall see later that this requirement in particular causes great difficulties in clinical medicine. Secondly, the classification must be exclusive, meaning that no observation must belong to more than one category, and thirdly the classification must be exhaustive, which means that all observations to be classified must belong to one of the categories. It is often possible to reduce a nominal scale to fewer classes. In the present example one might have used four classes (strings, woodwind, brass and percussion) or even two classes (‘stringed instruments’ and ‘other instruments’). A scale consisting of only two classes is called a binary scale.

Observations may be more refined. In 1806 the British admiral Sir Francis Beaufort constructed a scale for the measurement of wind force. The scale, which consists of 13 classes, is shown in Table 1.1. This is an example of an ordinal scale, and although it must have formed the basis for important decisions in the course of history, it has its limitations. We may take it for granted that a Force 10 is greater than, say, a Force 8, and that a Force 3 is greater than a Force 1, but we must not presuppose that the difference between Force 10 and Force 8 is the same as the difference between Force 3 and Force 1. That was revealed (as shown in Table 1.1) when it became possible to measure the wind force in m/s. To measure on an ordinal scale is like using an unevenly stretched elastic tape measure, and therefore it makes little sense calculating the ‘mean windforce’. An ordinal scale may also be reduced to a binary scale, which would be the case if we only distinguished between ‘windy weather’ and ‘calm weather’.
### Table 1.1 Beaufort scale of wind force. Specifications for use on land. The numbers in brackets indicate the wind force in m/s.

<table>
<thead>
<tr>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td><strong>Calm.</strong> Smoke rises vertically (0–0.2).</td>
</tr>
<tr>
<td>1</td>
<td><strong>Light air.</strong> Direction of wind shown by smoke drift, but not by wind wanes (0.3–1.5).</td>
</tr>
<tr>
<td>2</td>
<td><strong>Light breeze.</strong> Wind felt on face; leaves rustle; ordinary wane moved by wind (1.6–3.3).</td>
</tr>
<tr>
<td>3</td>
<td><strong>Gentle breeze.</strong> Leaves and small twists in constant motion; wind extends light flag (3.4–5.4).</td>
</tr>
<tr>
<td>4</td>
<td><strong>Moderate breeze.</strong> Raises dust and loose paper; small branches are moved (5.5–7.9).</td>
</tr>
<tr>
<td>5</td>
<td><strong>Fresh breeze.</strong> Small trees in leaf begin to sway; crested wavelets form on inland waters (8.0–10.7).</td>
</tr>
<tr>
<td>6</td>
<td><strong>Strong breeze.</strong> Large branches in motion; whistling heard in telegraph wires; umbrellas used with difficulty (10.8–13.8).</td>
</tr>
<tr>
<td>7</td>
<td><strong>Moderate gale.</strong> Whole trees in motion; inconvenience felt in walking against wind (13.9–17.1).</td>
</tr>
<tr>
<td>8</td>
<td><strong>Fresh gale.</strong> Breaks twigs off trees; generally impedes progress (17.2–20.7).</td>
</tr>
<tr>
<td>9</td>
<td><strong>Strong gale.</strong> Slight structural damage occurs (chimney pots and slates removed) (20.8–24.4).</td>
</tr>
<tr>
<td>10</td>
<td><strong>Whole gale.</strong> Seldom experienced inland; trees uprooted; considerable structural damage occurs (24.5–28.4).</td>
</tr>
<tr>
<td>11</td>
<td><strong>Storm.</strong> Very rarely experienced; accompanied by widespread damage (28.5–32.6).</td>
</tr>
<tr>
<td>12</td>
<td><strong>Hurricane.</strong> Disastrous results (&gt;32.6).</td>
</tr>
</tbody>
</table>

The *ranking* of data also provides measurements on an ordinal scale. School children may, for instance, be ranked from the top to the bottom of the form according to their proficiency, but once again one cannot assume that the difference between the proficiency of, say, numbers 5 and 6, is the same as that between numbers 6 and 7.

The *interval scale* represents the highest level of measurement. Weighing an object on a balance or the wind force in m/s may serve as examples. In these cases the scale is continuous and the interval, which is constant along the scale, is chosen to suit the precision of the measuring instrument. It may be 1 g for an ordinary letter balance and much less for an analytical balance. Interval scales may also be discontinuous (discrete). The number of patients in a ward may be 27 or 28, but not 27.5.

Usually, an interval scale is also a *ratio scale*. For instance, an object weighing 28 g is twice as heavy as an object weighing 14 g. Only measurements on a scale with an arbitrary zero point form an exception. Water having a temperature of 28 °C is not twice as warm as water having a temperature of 14 °C.

Measurements on an interval scale may be reduced to an ordinal or a binary scale. We may, for instance, measure objects in grams, but for some purposes we may confine ourselves to distinguishing between very heavy, heavy and light
Taking the history

Often the patient is not able to give a concise account of her complaints and the following example of a conversation between a patient and her doctor may serve as an example.

Doctor: ‘Good morning, Mrs N.N. What is the trouble?’
Patient: ‘I have had an uncomfortable feeling in my chest recently.’
D: ‘Have you got a pain?’
P: ‘I feel as though I cannot breathe.’
D: ‘When does this happen? When you exert yourself or when you are resting?’
P: ‘It only happens when I am working; when I do the cleaning I have to keep stopping to get my breath.’
D: ‘I remember you live on the second floor; how do you get on going up the stairs?’
P: ‘I take it slowly, but I have to wait on each landing.’
D: ‘How long has it been troubling you?’
P: ‘Ever since Christmas, but it got worse.’

This piece of conversation shows how the questioning progresses. The patient presents a vague complaint, ‘an uncomfortable feeling in the chest’, and it takes the doctor a couple of questions to be sure that the problem is dyspnoea on exertion. One may imagine that the doctor has in her mind a nominal scale of all possible symptoms, and that she is trying to refer her patient’s complaint to one of these named categories. Classification on a nominal scale, however, requires that all categories are well defined, and unfortunately most symptoms do not fulfil this requirement. All doctors may agree that the symptom angina pectoris suggests arteriosclerotic heart disease, but if participants in postgraduate medical courses are asked how they define the symptom, then the replies vary. Some say that it is necessary that the pain only occurs during exertion and others demand that the pain has a typical radiation. It always provokes heated discussion when doctors are asked to define common medical terms; there is immediate disagreement with an admixture of aggression, as everybody believes that her interpretation is the right one.
The patient in the example had dyspnoea, and that term is also ambiguous as it is used to describe both a subjective symptom (shortness of breath) and a physical sign (visibly laboured respiration). I shall return to the problem of definitions several times in this book, and I shall also point out that it does not matter so much which defining criteria are chosen, as long as there is mutual agreement. Authors of textbooks could do much to bring about such standardization.

It is best always to use neutral medical terms, i.e. terms that do not imply the cause of the complaint, and to avoid expressions such as ulcer dyspepsia, pleural pain and biliary colics. Such diagnostic data transformation is often found in medical notes and it may prejudice the clinician against other diagnostic possibilities. One should not attempt to make a diagnosis until the necessary data have been collected.

When the complaint has been named and classified its degree of severity must be assessed. In the example, the doctor extracted the information that the patient had to stop from time to time while doing her housework and that she must rest on each landing going upstairs. The severity of the symptoms can often be measured on an ordinal scale using expressions like mild, moderate and severe, but first it is necessary to define, as exactly as possible, the individual classes. We may, for instance, talk about mild dyspnoea if it does not limit normal activity, moderate dyspnoea if it limits but does not prevent normal activity, and severe dyspnoea if it prevents normal activity. Such an ordinal scale with predefined classes is often called a rating scale, and in the example the patient’s dyspnoea is rated as moderate. The severity of other symptoms as, for instance, abdominal pain in gastric ulcer patients may be assessed in exactly the same way: The patient notices the pain but continues working normally, her working capacity is reduced, or she stops working.

In these examples the severity was graded according to an ordinal scale with defined classes, but the severity of the symptoms in a particular patient may also be assessed over time by simple ranking. A patient may say: ‘I am feeling better than at the last visit, but I am not as well as I was this summer.’

A more refined method has been suggested for the measurement of pain intensity in, for instance, rheumatoid arthritis. The patients are asked to mark the degree of their complaint on a so-called visual analogue scale, i.e. a line, usually 100 mm long, ranging from ‘no pain’ to ‘pain as bad as pain can be’ (Fig. 1.2). Usually, however, this method cannot be recommended. Everybody knows what it means to suffer no pain, but the other end of the scale is left to the patient’s imagination, and in practice patients may be reluctant to commit themselves when they are shown the line.
The method also invites statistical problems. It is tempting to believe that the visual analogue scale is an interval scale and that it permits measurement of pain in mm, but that is not the case. It is only an ordinal scale, as it cannot be taken for granted that a difference of 10 mm at the lower end of the scale corresponds to the same difference in pain intensity as a difference of 10 mm at the higher end of the scale. Therefore, it makes little sense to calculate, for instance, mean and standard deviation, which presuppose an interval scale, and instead one should use ordinal scale statistics (with calculation of median and quantiles) as explained in Chapter 8. It has been shown that the grading of pain in rheumatoid arthritis, using a simple rating scale with four or five classes, is more sensitive than measurements on a visual analogue scale.\textsuperscript{25}

Finally, the symptom must be assessed chronologically. In the example the patient had suffered dyspnoea ‘since Christmas’ and this information is recorded as the number of months or weeks, i.e. on an interval scale. In the case of intermittent symptoms the chronological assessment also includes the duration of attacks and the intervals between them.

The physical examination

The physical examination, as it is performed today, originated in France at the beginning of the nineteenth century at which time the so-called Paris School of pathological anatomy was flourishing. Anatomical and clinical observations were correlated, and it became the task of the clinician, by means of the physical examination, to make those diagnoses that were known from autopsy studies.

Examination of the lungs and heart was developed at that time. Corvisart (1755–1821), Napoleon’s physician, made percussion a routine examination and Laënnec (1781–1826) developed auscultation by means of the stethoscope. It is said that once Laënnec had to examine an obese female patient, he could not feel the apex beat and in order not to apply his ear to her chest in an unseemly manner he rolled a sheaf of paper into a cylinder. Auscultation had its golden age before the introduction of radiology, electrocardiography and echography, but it is still of considerable importance. In general practice the
decision whether or not to treat a febrile patient with penicillin may depend on the result of the examination of the chest.

Unfortunately, the terminology that we use to describe our findings is chaotic. Anybody who spends a little time looking up the clinical signs in pneumonia, asthma and bronchitis in different textbooks will be surprised by the plethora of undefined terms. Perhaps we should confine ourselves to talking about coarse bubbling sounds (typically in chronic bronchitis), wheezing (in asthma), fine bubbling sounds (in pneumonia and congestive heart disease), and friction rubs (in pleurisy).

It is possible that computer technology will improve the situation by the introduction of educational programmes where the student sees the examination of the patient on the screen and at the same time hears what the physician hears in her stethoscope. In that way ostensive definitions, i.e. definition by demonstration of typical cases, may replace the usual verbal definitions.

In hospital practice, chest auscultation is part of the routine physical examination, but other clinical examinations are only done on those patients who present certain clinical pictures, e.g. palpation of the temporal artery in patients with pyrexia of unknown origin or examination for ruptured ligaments after a knee trauma. Such examinations represent diagnostic tests that are carried out on the suspicion of a particular diagnosis.

Paraclinical findings

Typically, paraclinical findings are the results of all those tests and observations that are made in the ‘paraclinical’ departments, i.e. departments of pathology, clinical chemistry, radiology etc. (but they may also be said to include those simple laboratory tests which are done by the clinicians themselves). They serve a clinical purpose and, therefore, they are included when in this book I use the term clinical data or talk about the patient’s clinical picture.

Radiological, scintigraphic and histological findings are examples of descriptive, paraclinical data. They must, just like symptoms and physical signs, be classified by means of a nominal scale with well-defined classes (e.g. different types of colonic polyps), and they may be graded on an ordinal scale (e.g. slight, moderate or severe dysplasia).

Paraclinical observations are usually considered more objective, and therefore more reliable than those observations that are made at the bedside. This may often be true, but clinicians should be aware of the fact that, for instance, X-ray reports may be biased by the information on the request form, in which
case the paraclinical observation may just help to sustain a wrong diagnosis. In order to solve this problem, the suggestion has been made that the paraclinical diagnostician should receive no information about the patient’s clinical picture, but that is of course quite unrealistic. A radiologist, for instance, must know which clinical problem she is expected to solve, just as the physician doing a clinical examination will take into account whether the patient is febrile, has abdominal pain or presents other symptoms. However, unnecessary diagnostic data transformation is to be avoided. The radiologist who sees a consolidation in one lung should strive at a neutral description of this finding, and she should not write ‘pneumonia’ if it is noted on the request form that the patient has a high temperature, but something else if the temperature is normal. The radiologist’s expert advice is, of course, needed when the clinician considers various diagnostic possibilities, but ideally the synthesis of the radiological findings and the rest of the information about the patient should take place after the initial description of the films, for instance, at a joint X-ray conference. In clinical research, however, it is essential that the X-ray films are read by a radiologist who has no knowledge of the other components of the clinical picture.

Histological reports present a special problem as the pathoanatomical diagnosis frequently also represents the patient’s diagnosis, simply because the disease classification to a large extent is anatomically orientated. It is, however, necessary to be aware of the fact that doctors representing different specialties sometimes attach very different meanings to the same terms. Gastritis is a good example. Particularly in the past, general practitioners have often used this term as a diagnostic label for patients with upper abdominal dyspepsia. Gastroenterologists use the term to denote certain endoscopic findings, and pathologists only talk about gastritis if histological sections of a biopsy show inflammatory or atrophic changes.

A different example is provided by the words ‘acute’ and ‘chronic’. To the clinician ‘acute’ means ‘of sudden onset’ and ‘chronic’ is used in the sense ‘persistent’, whereas to the pathologist ‘acute’ is synonymous with ‘infiltrated with polymorphs’ and ‘chronic’ denotes ‘infiltrated with round cells’. A rectal biopsy from a patient with ulcerative colitis usually presents ‘acute inflammatory changes’ in spite of the fact that the patient may have had the disease for many years.

The concentrations of different substances in the patient’s blood are typical examples of quantitative paraclinical data, representing measurements on an interval scale. Such results are usually less ambiguous than the descriptive data, but the clinician should remember that methods of analysis vary and that results from different laboratories may not be comparable.
Global assessments

Up to now I have considered the assessment of the individual components of the clinical picture, but often clinical decisions are based not only on the presence or absence of a particular symptom or sign, but also on a global assessment of the patient’s condition. The general practitioner who visits a child with a high temperature may not be able to make a specific diagnosis and then she will rely on her global evaluation of the child's condition. A child who ‘looks ill’ may be admitted to hospital at once, whereas the child who seems relatively unaffected will be observed at home.

Global assessments are also essential when clinicians assess the effect of their treatments, both in the daily routine and in clinical research. Therefore, it is unfortunate when papers in medical journals concerning the effect of cancer chemotherapy only report the effect on survival. It is, of course, important to know the average number of days that the patients survived when they received different treatments, but this information ought to be supplemented with global assessments of their condition during the course of the disease. It must be reported how they felt, what they were able to do, and how the harms of the treatments affected them.

It may be useful for research purposes, but also in daily clinical practice, to formalize such global assessments by calculating a clinical index and for that purpose one may use simple rating scales or more complex scoring systems. The rating scales are constructed much like the Beaufort scale (see Table 1.1). The number of classes may vary, but, if a simple five-point scale is used, the patients may record at the end of the treatment period (provided it is so short that they can remember how their condition was before treatment) whether they feel much better, a little better, the same, a little worse or much worse. The ratings may also be observer ratings, i.e. ratings done by the clinician, but in both cases it is important, as Beaufort did, to define the classes as well as possible.

The New York Heart Association (NYHA) classification of the functional capacity of patients with cardiac incompetence is a good example of a rating scale. It uses the following criteria (abbreviated):

1. No limitation of physical activity;
2. Ordinary physical activity results in fatigue, palpitation or dyspnoea;
3. Less than ordinary activity causes fatigue, palpitation or dyspnoea;
4. Symptoms at rest.
In other cases the components of the global assessment cannot be measured on the same scale, and then one must rate each component independently and afterwards calculate a total score. A well-known example is the Apgar score for assessing the state of the newborn. Here the heart rate, respiratory effort, muscle tone, response to stimuli and colour are rated independently as 0, 1 or 2, and the total score is calculated by simple addition of the five ratings.27 Other scoring systems, however, are much more complicated. Some anaesthesiologists, for instance, use the APACHE score to assess the prognosis of critically ill patients, which includes both a clinical assessment and a large number of laboratory tests (e.g. electrolytes, arterial pH and the white blood cell count).28

The calculation of clinical indices may be useful, but it is a condition that they make sense from a clinical point of view, and that is not always the case. More than 70 different methods for assessing the treatment effect in rheumatoid arthritis have been used, and some of them make little sense.29 Previously, Ritchie’s index was often calculated. Here the doctor palpates a number of joints, records the pain response on a rating scale from 0 to 3, adds up the results and calculates the total score. In other studies the doctor simply counts the number of tender or swollen joints. When these methods are used, two patients, one with a severe pain in one or two joints and another with a slight pain in a larger number of joints, may receive the same score, although, obviously, the former patient is much more incapacitated than the latter. In trials of schizophrenia, more than 600 different scales have been used30 and, obviously, they cannot all have been meaningful.

Global assessments made by the patients themselves by means of a visual analogue scale are also of little use. The end points of the scale are defined as ‘the best imaginable’ and the ‘worst imaginable’ condition which means that their definition is left to the imagination of the patient. In most cases it is better to use simple rating scales with few classes.25

No matter which method is used, it must be remembered that we are dealing with measurements on an ordinal scale, and, although it is often done, it makes little sense to calculate the mean and the standard deviation (Chapter 8). Complex clinical indices are sometimes referred to as quality of life measurements, but this term should not be used in this context (Chapter 7).