COMPUTER ANALYSIS OF SYMPTOM PATTERNS AS A METHOD OF SCREENING PATIENTS AT SPECIAL RISK OF HYPOTHYROIDISM

D. J. P. BARKER, Ph.D., B.Sc., M.R.C.P.
AND
J. M. BISHOP, D.Sc., M.D., F.R.C.P.

Department of Medicine, University of Birmingham, Queen Elizabeth Hospital, Birmingham, 15

Patients in the Birmingham region who are treated with radio-iodine are kept under continued surveillance by means of a follow-up system, which effects periodic screening for hypothyroidism. In this follow-up system, which is described elsewhere (Barker and Bishop, 1969), the serum protein bound iodine (PBI) is measured at intervals of 21 months. Nine months before serum PBI measurement is due a postal questionnaire is sent to each patient. It comprises nine questions, each of which elicits a symptom of hypothyroidism and must be answered 'Yes' or 'No'. If at any time replies to the questionnaire, or the level of the serum PBI, suggest hypothyroidism, the patient is referred to the out-patient clinic for full medical assessment.

METHOD OF ANALYSIS OF THE QUESTIONNAIRE

After a preliminary study, completed questionnaires were received from 338 patients, of whom 65 were subsequently shown to be hypothyroid. The decision that a patient was hypothyroid, and required thyroxine, was made independently of the questionnaire replies; so that it was possible to devise criteria from the 338 questionnaires which would enable questionnaires received subsequently to be classified as euthyroid or hypothyroid.

The nine symptoms elicited by the questionnaire were malaise, cold intolerance, loss of appetite, loss of energy, increase in weight, change in the voice, fullness of the face, dryness or roughness of the skin, and untidiness of the hair. None of these symptoms is specific to hypothyroidism. They occur in patients who have other disorders or who are ageing. Since almost all patients are older than 50 at the time of treatment with iodine-131 (\(^{131}I\)) the subsequent development of such symptoms as cold intolerance in euthyroid patients is not uncommon. Questionnaires cannot therefore be identified as coming from hypothyroid patients by the presence of certain specific symptoms. It is necessary to adopt a more complex method of classification, the criteria for which will depend upon the presence or absence of nine symptoms in any individual.

Theoretically there are \((2^n - 2)\) different criteria which can be used to define a subgroup from a group of individuals, each of whom may or may not have any one of \(n\) symptoms. Thus within a group of patients in whom two symptoms \(A\) and \(B\) may be either present or absent 14 different subgroups of patients can be defined. These subgroups are listed in Table I where, for example, the sub-

<table>
<thead>
<tr>
<th>(A\B)</th>
<th>(A'B)</th>
<th>(AB)</th>
<th>(AB\B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(AB)</td>
<td>(AB\B)</td>
<td>(A'B)</td>
<td>(AB)</td>
</tr>
<tr>
<td>(AB)</td>
<td>(AB\B)</td>
<td>(A'B)</td>
<td>(AB)</td>
</tr>
<tr>
<td>(AB)</td>
<td>(AB\B)</td>
<td>(A'B)</td>
<td>(AB)</td>
</tr>
<tr>
<td>(AB)</td>
<td>(AB\B)</td>
<td>(A'B)</td>
<td>(AB)</td>
</tr>
<tr>
<td>(AB)</td>
<td>(AB\B)</td>
<td>(A'B)</td>
<td>(AB)</td>
</tr>
</tbody>
</table>

The symbol \(\overline{A}\) indicates absence of symptom \(A\)

The group \(AB\), \(\overline{A}\B\) includes two types of patients, who are respectively coded as \(AB\) (symptom \(A\) present, symptom \(B\) absent) or \(\overline{A}\B\) (symptom \(A\) absent, symptom \(B\) present). The subgroup \(AB\), \(\overline{A}\B\) therefore includes patients who have either one or other of the two symptoms, but excludes patients who have both symptoms (\(AB\)) or neither of the symptoms (\(\overline{A}\B\)). Likewise, the subgroup \(AB\), \(\overline{A}\B\) includes both types of patient who have symptom \(A\) (\(AB\) and \(\overline{A}\B\)) and those who have neither symptom (\(\overline{A}\B\)), but excludes those in whom \(B\) is present but \(A\) is absent (\(\overline{A}\B\)).
Where \( n \), the number of symptoms, is greater than 4 the number of possible subgroups exceeds
10\(^n\), and any systemic investigation becomes impossible. The postal questionnaire elicits the
presence or absence of nine symptoms in each individual, and the selection of one subgroup of
individuals who will be classified as 'suspected hypothyroid' cannot be carried out by systematic
testing of every possible subgroup, to determine which contains the proportion of hypothyroid
patients that is optimal for screening purposes. In order to devise a method for classifying replies to
the questionnaire it is necessary to make some initial, partly arbitrary simplification.

Considering the natural history of hypothyroidism, it seems a reasonable assumption that,
following an abrupt or gradual decline in thyroid hormone production, the evolution of symptoms
will proceed along a number of definable pathways. If this is so, certain combinations of symptoms will
characterize certain forms or stages of the disorder, and questionnaires completed by hypothyroid
patients may be recognized by the presence of these symptom-combinations.

When, during follow-up, a completed questionnaire is received, each symptom-combination shown is considered. The number of symptom-combinations equals \( 2^n - 1 \), where \( n \) is the number of symptoms; and a patient may therefore have up to 511 symptom combinations, with from one to nine symptoms in each combination. For each combination calculation is made of the percentage \( x \) of hypothyroid patients among all patients previously observed to have that particular combination. Thus a patient with three symptoms C, D, and E shows seven symptom-combinations, C with D, C with E, D with E, C with D with E, C alone, D alone, and E alone, for which the seven values of \( x \) might be 20%, 25%, 50%, 35%, 10%, 15%, 10%. In this case the combination of D with E \( (x = 50\%) \) is the symptom-combination most specific to hypothyroidism since it is associated with the highest value of \( x \). In the follow-up system a patient is classified as 'suspected hypothyroid' if the highest value of \( x \) exceeds a predetermined limit.

Table II shows the highest values of \( x \) for each of the 338 questionnaires received in the preliminary study. If patients are classified as hypothyroid when \( x \) is greater than 39%, then 46 (71%) hypothyroid patients are correctly classified, and 80 (29%) euthyroid incorrectly classified. Selection of this value of \( x \) to separate 'suspected hypothyroid' from euthyroid patients is, as in all screening procedures, determined by the number of

<table>
<thead>
<tr>
<th>Highest Value of ( x )</th>
<th>No. of Euthyroid Patients</th>
<th>No. of Hypothyroid Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 10</td>
<td>37</td>
<td>3</td>
</tr>
<tr>
<td>10-20</td>
<td>76</td>
<td>5</td>
</tr>
<tr>
<td>20-30</td>
<td>80</td>
<td>11</td>
</tr>
<tr>
<td>30-40</td>
<td>39</td>
<td>15</td>
</tr>
<tr>
<td>40-50</td>
<td>33</td>
<td>23</td>
</tr>
<tr>
<td>50-60</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>60 and over</td>
<td>273</td>
<td>65</td>
</tr>
</tbody>
</table>

*See text for definition of \( x \). If a symptom combination has been recorded in less than 10 patients no value for \( x \) is calculated.

Analysis of the questionnaires is carried out by computer, which calculates the highest value of \( x \) and, if this exceeds 39%, prints out a directive that a serum PBI value is required at once. The patient is referred to the out-patient clinic if the PBI is abnormal, otherwise follow-up is continued by another questionnaire sent in one year's time. If the highest value of \( x \) exceeds 49%, the patient is referred to the out-patient clinic irrespective of the PBI result. Of such patients 43% are hypothyroid (Table II), and for this reason it seems desirable that the diagnosis of hypothyroidism should be excluded after full medical assessment in addition to a single measurement of the serum PBI.

The values of \( x \) are calculated by reference to data which are stored on magnetic disc, and which are derived from questionnaires returned by patients whose thyroid state has been subsequently determined. This store comprises two arrays of 511 units, each corresponding to a particular combination of symptoms. Data from hypothyroid patients are stored in one array, and those from euthyroid patients in the other. The initial data in the store related to the 338 patients in the preliminary study, but as more patients come under review more completed questionnaires are added to the store. If a questionnaire is followed by a normal serum PBI nine months later (or immediately if the questionnaire is classified as 'suspected hypothyroid'), the questionnaire is stored in the euthyroid array. If a questionnaire is followed immediately by an abnormal serum PBI, it is classified as euthyroid or hypothyroid, according to the decision made when the patient attends the out-patient clinic. If a questionnaire is
followed by an abnormal serum PBI nine months later, it cannot be stored, as the patient’s thyroid state at the time of completion of the questionnaire is unknown. In this case another questionnaire is sent to the patient at the time of attendance at the clinic. The stages in the analysis and storage of questionnaires are shown in the Figure.

**DISCUSSION**

Computer analysis of patterns of replies to a postal questionnaire is an inexpensive screening procedure which requires only occasional medical supervision. It causes no inconvenience to the patients and therefore does not deter them from continuing to cooperate with the follow-up. The results of the preliminary study (Table II) show that a proportion of hypothyroid patients will be misclassified as euthyroid, and it is clearly necessary to use postal questionnaires in association with a more accurate screening test such as the serum PBI. If the postal questionnaires were not used in the follow-up system this test would have to be performed more frequently, with the need for additional hospital visits by the patient and with an increase in the cost of the system. In the preliminary study one hypothyroid patient admitted to no symptoms either in reply to the questionnaire or at subsequent interview, and it therefore seems that no screening procedure which depends on symptoms can identify all of those with the disorder.

The method used to classify the questionnaires is an empirical one, selected because it conformed with certain assumptions which can be made about the natural history of hypothyroidism, and because it was subsequently shown to give a sufficiently small number of ‘false positives’ and ‘false negatives.’ The method requires no assumptions about the dependence or independence of symptoms, whereas in statistical methods based on Bayes’s theorem, which have been widely used in automated diagnostic systems, it is necessary to assume that symptoms are independent. For the reasons already given, it is impossible to explore all the empirical methods which could be devised to classify the questionnaires, and it is likely that some of these methods would give more discriminating results than the method used. It is, however, of interest that by the use of one obvious alternative method, in which patients are classified as hypothyroid when more than a certain number of symptoms are present, it is not possible to obtain such good discrimination between hypothyroid and euthyroid patients. If patients in the preliminary study are classified as hypothyroid when four or more symptoms are present, 44 (67%) hypothyroid patients are correctly classified but 112 (41%) euthyroid patients are incorrectly classified.

During analysis of a questionnaire the values of $x$ are calculated by reference to a store of data, which is continuously supplemented as the follow-up progresses. This enables each questionnaire to be classified after reference to all data collected up to that time. Periodically it will be necessary to review the threshold value of $x$ which divides euthyroid and hypothyroid patients, since the proportion of hypothyroid patients in the group being screened will change. After the follow-up system has been in operation for some years it may
D. J. P. BARKER AND J. M. BISHOP

become possible to classify a questionnaire by comparison with previous replies from the same patient, rather than by reference to the frequencies of symptom-combinations among all patients in the group.

A group of screening procedures which forms part of a periodic health examination in San Francisco includes a questionnaire eliciting symptoms of bronchial asthma (Colhen, Rubin, Neyman, Dantzig, Baer, and Siegelaub, 1964). There are six questions and therefore $2^6$ or 64 ways of completing the questionnaire. The statistical method used to classify the questionnaire is based upon a method described by Neyman (1950). For each of these 64 types of reply likelihood ratios are calculated by relating the proportion of asthmatic patients who gave the reply in a preliminary study to the proportion of non-asthmatic patients. In the screening programme patients are classified as asthmatic if they give a reply for which the likelihood ratio exceeds a predetermined limit. This method seems to have no advantage over the one used in the present study. Mathematically, calculation of likelihood ratios is unnecessarily complex, since the same classification would be obtained by relating the numbers of asthmatic and non-asthmatic patients giving each reply. Furthermore, it does not seem that this empirical classification needs a formal basis in theories for testing statistical hypotheses. A method based on consideration of each type of reply to the questionnaire is of limited use, since as $n$ the number of questions, increases, $2^n$, becomes very large; and it would require a very extensive preliminary study to establish the frequencies of the disorder associated with more than a small proportion of the total number of types of reply.

The results of analysis of replies to postal questionnaires cannot be compared with other published results of analyses of data on hypothyroidism, since these have depended on information elicited from interviews and clinical examination. Billewicz, Chapman, Crooks, Day, Gossage, Wayne, and Young (1969) have devised a diagnostic index which effectively divides clinic patients suspected of hypothyroidism into euthyroid, hypothyroid and doubtful categories. A score is allocated to each symptom or sign when present or absent, and discrimination depends on the total score of each patient. The scores are determined from comparison of the frequency of each symptom and sign in hypothyroid and euthyroid cases. However, the method of classification which we have described seems more appropriate to the theoretical requirements of the data than one derived from comparisons of frequencies of individual symptoms. If, during the evolution of hypothyroidism in patients, symptom A develops at the same time as B, and C develops with D, then a patient who has A and C may not be correctly classified as euthyroid if the method of analysis depends only on consideration of frequencies of individual symptoms. The symptom of dryness of the hair was not included in the diagnostic index because it was found to occur more frequently among euthyroid than hypothyroid cases. However, it is possible that, in combination with one or more other symptoms, dryness of the hair is characteristic of hypothyroidism.

With even a small number of symptoms the numbers of combinations which may occur is very large; and it has hitherto been impossible to obtain quantitative statements about the significance of particular combinations of symptoms, or of signs or laboratory results. However, because computers have the facility of storing large arrays of data, patterns of clinical attributes shown by patients can now be recorded. This offers a new method of studying clinical situations.

**SUMMARY**

A postal questionnaire which elicits symptoms is used as part of a screening system for patients at special risk of hypothyroidism. Completed questionnaires are analysed by computer, and those returned by hypothyroid patients are recognized by the presence of particular combinations of symptoms associated with a high frequency of the disorder. It is suggested that similar analyses of combinations of symptoms, signs or laboratory results will provide information that will assist clinical decisions.

This study was facilitated by a grant from the Endowment Research Fund of the United Birmingham Hospitals.

**REFERENCES**


Computer analysis of symptom patterns as a method of screening patients at special risk of hypothyroidism.

D J Barker and J M Bishop

doi: 10.1136/jech.24.4.193