J Med Genet 1990; 27: 446–450

EL BUSCA and the value of signals in the diagnosis of dysmorphic syndromes: good and bad handles in computer assisted differential diagnosis

L J Salgado, J S Lopez-Camelo, E E Castilla

Abstract

A computer system for the assistance of syndrome diagnosis in dysmorphology (EL BUSCA) was developed, and used to test the mechanics of the diagnostic process. EL BUSCA has a reference file (REF) with 200 syndromes, expressed in 175 signals. Signals have a weight value resulting from the difference between the number of syndromes including that sign and the total number of syndromes in the REF. A mean signal weight was calculated for each syndrome. The system was tested with 200 published cases (CASES), representing 82 different syndromes. Each consultation (CONS) entered up to 15 patient signals. The system then selected syndromes having three or more of those signals. 'Present' (REF+CASE), 'Absent' (REF only), and 'Additional' (CASE only) signals, as well as the score given by the sum of the weights of 'present' signals, were displayed for each suggested diagnosis. A consultation was successful (positive answer) if the correct diagnosis appeared among the first 12 ranked. EL BUSCA gave a positive answer in 82% of the 200 test consultations.

Linear regression, with ranking of the correct diagnosis among the answers as the dependent variable, was used for the analysis of the following results. For the REF, no relationship was found for either the number or the mean weight of the signals with the ranking of the correct diagnosis. For the CASES, there was a linear relationship between the number of signals of each consultation and the ranking of the correct diagnosis, indicating that the larger the number of signals consulted, the lower the ranking of the correct diagnosis. No effect was seen for the mean weight of consulted signals. For the CONS, the number and the weight of present and additional signals did influence the correct diagnosis, while no effect was seen for absent ones.

EL BUSCA will not be developed further as a diagnostic tool since the availability of the London Dysmorphology Database (LDDB) system widely satisfies the need in clinical dysmorphology. Further improvement in the LDDB list of signals would certainly increase its diagnostic power.

Hospital Lucio Melendez, Adrogue, Argentina.
L J Salgado

IMBICE: Instituto Multidisciplinario de Biologia Celular, La Plata, Argentina.
J S Lopez-Camelo

FIOCRUZ: Dept Genetica, Instituto Oswaldo Cruz, Rio de Janeiro, Brazil.
E E Castilla

Correspondence to Dr Castilla, ECLAMC/Genetica/FIOCRUZ, CP 926, 20010 Rio de Janeiro, Brazil.

Received for publication 12 October 1989.
Accepted for publication 5 January 1990.

The diagnostic process is often difficult in clinical dysmorphology because of the large number of syndromes described, the low prevalence of each of them, and the constant flow of new information in this area.

The introduction of computing techniques in medical diagnosis by INTERNIST-I[1] helped in the process of differential diagnosis in general. Several other diagnostic systems were then created, some of them specialising in dysmorphology, such as GENDIAG,[2] SYNDROC,[3 4] and the LDDB,[5] among others. The design and use of these systems brought a better understanding of concepts intuitively used in clinical diagnosis, including predictive value, sensitivity, and significance of signals and symptoms.

In this paper we describe the development of a new computing system, EL BUSCA, to assist in the diagnosis of dysmorphic syndromes. Special emphasis is given to the discussion of the theoretical matters involved.

Material and methods

THE REFERENCE FILE

A reference file, based on D W Smith's classical atlas,[6] was created, including 200 syndromes and a list of all possible signals reported for each one, with no limitation of maximum number.

In order to obtain a ranking of the suggested

diagnosis, each sign was weighted by the difference between the number of syndromes including that sign and the total number of syndromes in the reference file. Therefore, the weight of each sign is inversely related to its frequency among the syndromes in the reference file, and expresses its predictive value. A mean weight was also calculated for each selected syndrome by adding the weight of all its signals, divided by the number of signals.

In an effort to increase their diagnostic value, the list of 158 signals present in Smith's atlas[6] was modified to a final number of 175 signals. Some of them were subdivided: deformed occiput into flat and prominent; abnormal philtrum into short, long, and other; slanted palpebral fissures into mongoloid and antimongoloid, etc. New signals were also introduced, mainly easily observed and *x* ray signals such as kyphosis, lordosis, hemivertebrae, osteoporosis, etc.

Since the main objective was to test a diagnostic system instead of developing a diagnostic tool in its final form, the list of malformation patterns present in Smith's atlas[6] was modified in several ways. Some patterns were deleted, including three large diagnostic categories: chromosome anomalies and skeletal and ectodermal dysplasias. The reasons included too few defects involved (polysyndactyly) or too obvious diagnosis (anencephaly sequence). Others were added, including the most recently described syndromes, as long as enough cases had already been published to be reasonably accepted as nosological entities (for example, acrocallosal, Proteus, Golabi-Rosen syndromes, etc).

TEST CASES

In order to test the system, 200 cases, representing 82 different dysmorphic syndromes, were selected from mainstream specialised journals. In this way, the correctness of the given diagnosis was validated by the authors, the referees, and the editors. Reported cases were accepted when three or more signals present in the reference file were mentioned; superficially described cases were discarded. Four of the 82 syndromes had six test consultations each, while the remaining had fewer than that.

Only case reports were considered; reviews, in which signals are usually simplified and tabulated, were avoided. In family reports, only the proband was included because his/her description is usually given in more detail. In cases with long term follow up, with new signals and symptoms appearing later in life and some early ones perhaps disappearing, all signals were recorded, no matter how transient they might be.

THE CONSULTATION

A consultation consisted of entering the signals present in the patient, up to a maximum of 15. The system searches and selects those syndromes having three or more of the patient's signals. Both numbers, 3 and 15, were arbitrarily set because they were a priori considered as appropriate for the number of signals usually described for a given patient.

For each suggested diagnosis, the display gives the following information. (1) **Present** signals, found in both patient and reference syndrome. (2) **Absent** signals, only found in the reference syndrome. (3) **Additional** signals, found only in the patient. This information was meant as a guideline for further physical examinations. The diagnosis itself, however, was made by the consultant after evaluating the clinical significance of the additional signals listed, as well as other pieces of information not included in the consultation, such as family history, age, sex, etc.

The suggested diagnoses are ranked according to an absolute value, resulting from adding the weights of 'present' signals (shared by the patient and the reference file), and displayed by decreasing score, up to a maximum of 12.

A given consultation was considered successful (positive answer) when the correct diagnosis appeared among the first 12 in the ranking. Results were analysed by linear regression.

Results

EL BUSCA gave a positive answer in 82% (165/200) of the test consultations, occupying the first two positions in 59% of the consultations (118/200).

Three series of data were analysed in the search for characteristics associated with the outcome of the consultations. Those corresponding to (1) the 82 correct diagnoses in the reference file (REF), corresponding to the 200 test cases consulted; (2) the 200 test cases (CASES), and (3) the consultation (CONS), that is, the relationship of each consulted case with its correct diagnosis as defined in the reference file: the **Present, Absent,** and **Additional** signals of those cases with a positive answer.

These three series of data are simple listings of signals: (1) signals of each consulted syndrome, (2) signals of each consulted case, (3) signals within each present, absent, and additional category. For each of the three data groups, the number and weight of the signals involved were analysed in connection with the ranking given to the correct diagnosis.

In all analyses, the ranking was the dependent variable, with number and weight of signals as the independent ones. Signal mean weight was considered for the test cases and reference file data sets, while absolute weight was taken for the consultation's present, absent, and additional signals.

ANALYSIS OF REFERENCE FILE

This set of analyses tried to identify those character-

istics of each syndrome influencing its ease of diagnosis. No linear relationship could be found for either the number ($t=-1\cdot1$, $p>0\cdot05$) or the mean weight ($t=-0\cdot6$, $p>0\cdot05$) of the signals with the ranking post of the correct answer.

ANALYSIS OF TEST CASES

A linear relationship ($t=2\cdot0$, $p<0\cdot05$) was found between the number of signals of each consultation and the ranking obtained for the correct diagnosis, indicating that the larger the number of signals consulted, the worse the diagnosis obtained. The mean number of signals per consultation was 10·3 (SE=0·2). When the mean weight of the signals for each consultation was related to the ranking obtained for the correct diagnosis, a non-significant association was seen ($t=-1\cdot4$, $p>0\cdot05$).

A linear relationship ($t=2\cdot0$; $p<0\cdot05$) was found between the number of signals of each consultation and the ranking obtained for the correct diagnosis, indicating that the larger the number of signals consulted, the worse the diagnosis obtained. The mean number of signals per consultation was 10·3 (SE 0·2). When the mean weight of the signals for each consultation was related to the ranking obtained for the correct diagnosis, a non-significant association was seen ($t=-1\cdot4$, $p>0\cdot05$).

ANALYSIS OF THE CONSULTATION

The number of signals

When the number of present, absent, and additional signals was analysed, as independent variables, against the ranking post of the correct diagnosis in each of the 165 consultations with a positive answer, the relationship was evident ($F=125\cdot9$, $p<0\cdot01$). The influencing factors, in this respect, were both present ($F=164\cdot7$, $p<0\cdot01$, $r2=28\cdot7\%$) and additional ($F=211\cdot9$, $p<0\cdot01$, $r2=51\cdot0\%$) signals, while no influence was proven for the absent ones ($F=1\cdot0$, $p>0\cdot05$, $r2=0\cdot7\%$).

The weight of signals

The weight of present, absent, and additional signals also proved to influence the ranking post of the correct diagnosis ($F=132\cdot3$, $p<0\cdot01$). As seen for the number of signals, the significant factors were the present ($F=195\cdot4$, $p<0\cdot01$, $r2=33\cdot0\%$) and additional ($F=200\cdot9$, $p<0\cdot01$, $r2=51\cdot0\%$) signals, but not the absent ones ($F=0\cdot4$, $p>0\cdot05$, $r2=0\cdot5\%$).

Discussion

EL BUSCA

Better delineated syndromes were supposed to have more signals in the reference file, being better known and easily recognised. Similarly, syndromes with a large proportion of specific signals should have a higher mean weight, being easily recognised too. However, neither the number nor the mean weight of the signals of syndromes in the reference file was related to the outcome of consultations. Even when these suppositions were true in some cases, the following opposite situation also occurred frequently. New syndromes, still with only a few published cases and therefore describing the most severe end of the clinical spectrum, can be identified by only a few signals simply because few signals are available for that diagnosis. Furthermore, in some cases, the test case could have been one of the few considered to delineate the syndrome in the reference file.

The observation that the lower the number of signals entered in a given consultation, the higher the ranking for the correct diagnosis, seems, at first sight, to go against clinical common sense. However, several factors may be claimed to explain this result. A published diagnosis is based, most of the time, on a selection of the patient's signals, reflecting the author's subjectivity; he already has a diagnosis in mind which he tries to prove. Consequently, common minor defects will have less chance of being mentioned when associated with three or four well defined, major defects. Next, those syndromes that can be defined with few signals usually have few competitors in the process of differential diagnosis. The shorter the list of alternatives, the better the chances of ranking high on it. This probably reflects the fact that fewer words are needed to express certainty than doubt.

Our results did not prove the underlying hypothesis that patients with more specific signals would have more evident diagnoses, ranking high during testing consultations. This could reflect the lack of predictive value of the weight of the signals in EL BUSCA. Such limitation does not invalidate the rationale applied, as it derives from an incomplete reference file.

While both the number and the weight of the 'present' signals are associated with the correct diagnosis, weight seems to have a larger influence than number. This difference would be greater if the reference file was improved, making the signals more specific, that is, heavier.

In contrast to internal medicine,[1] the sensitivity (its frequency among the affected) of each signal in each syndrome is generally unknown in dysmorphology, impairing the evaluation of the 'absent' signals.

The determination coefficient ($r2=51\cdot0\%$) of the 'additional' signals to the ranking of the correct diagnosis, being higher than that of the 'present' signals ($r2=33\cdot0\%$), may indicate that the presence in a patient of a signal not expected for a given syndrome bears more against it being the correct diagnosis than the presence of an expected signal favours it. This is a consequence of the principle of parsimony, by which a patient must have only one disease at a time, all

manifestations resulting from a single cause. This principle is more strictly followed in aetiopathogenetically well understood fields, such as internal medicine, than in dysmorphology, where the diagnosis of patients and the delineation of syndromes are still investigated more or less simultaneously. Following the theory of taxonomy, by which an unidentified object is placed into its corresponding class ('determination') according to a pre-established system ('classification'),[7] at the present stage of knowledge in dysmorphology, the process of 'determination' acts backwards, influencing that of 'classification'.

THE DIAGNOSTIC PROCESS IN DYSMORPHOLOGY

A diagnosis may be heuristic knowledge, resulting from logical reasoning, only when the cause (aetiology) and the affected process (pathogenesis) are known.[8] Since this is seldom the case in dysmorphology, other ways of recognition have to be used. The Bayesian and even the pseudo-Bayesian methods are not suitable because population frequencies of signals and syndromes are mostly unknown. Therefore, we are forced to be Gestaltic, approaching recognition by likeness between patients.

On the other hand, today's fourth generation computers only allow for a linear rationale, adequate for heuristics but not for Gestalt. Until artificially intelligent systems are developed, computer assisted diagnosis will necessarily have to rely upon the operator's gestaltic knowledge in the first step of carefully selecting a few 'good handle' signals,[5] as well as in the last step of deciding the final diagnosis from the proposed listing of possibilities.

A 'good handle' is, for the LDDB,[5] a signal having two characteristics, to be clearly pathological and to be rare among syndromes. The first one, which makes it relevant to the process of diagnosis, was solved by INTERNIST-1[1] with a significance value given to each signal, and by GENDIAG,[2] by weighting the signals involved. However, EL BUSCA did not consider the frequency of each signal in the general population. Hence, signals which are common in the normal population but rare among syndromes obtained an undeservedly high weight. Such was the case, for instance, for 'prominent supraorbital ridges', being weighted 198 because present in only two of the 200 reference syndromes. Even when differential diagnosis is made among syndromes, excluding the state of normality as an alternative, the high frequency of a signal in the general population diminishes its diagnostic relevance because its presence in the patient may be coincidental.

EL BUSCA based the weight of signals on the second condition for a 'good handle', its rarity. The probability of having a given illness in the presence of a given signal is the property that more clearly connects a consulted case with a diagnosis because here the reasoning flows in the same direction as the problem proposed. This specificity would be absolute in the case of a signal being present in only one syndrome, that is, pathognomonic, an extreme situation almost non-existent in clinical practice. This rarity among syndromes gives the signal its predictive value, or specificity, which is called 'evoking strength' by INTERNIST-1[1] and it is expressed by coordinates M1 and M2 in the original version of SYNDROC.[2]

The methodology for diagnostic investigation used by the LDDB[5] gives the elected signal (handle) an absolute value, since only those syndromes having it are included in the proposed list. For this reason, a series of several consultations are required, combining different sets of signals. Even when the authors state that the LDDB assumes no weight for the signals, the list of proposed syndromes given in each consultation does actually constitute what in taxonomy is named a monothetic group, that is, where all its members (syndromes) share a given set of attributes (signals), which is a necessary and sufficient condition to be a member of the group. As a matter of fact, most well known syndromes are polythetic groups, with their members sharing a number of attributes, the only one shared by all members being the attribute of belonging to the group.[7]

The process of diagnosis may be seen as the 'discover' of a hidden reality, the syndrome, manifested by a language of signals. As in any language, some degree of generalisation is always needed, it being an intermediate reality between a subject (the clinician) and an object (the patient). For instance, the description of the facies in a multiply malformed patient may range from the generic 'peculiar facies' to the detailed description of several features, to the shape of every relief of the pinnae. Both levels will be of little help because they depart from the intermediate point between subject and object required, for optimal performance, in the language of signals. This midpoint varies according to the observer, the material to be discriminated, and the reason for discrimination. For instance, the same nose will have a different optimal description for plastic surgeons, ENT doctors, or geneticists.

EL BUSCA has never been used as a diagnostic tool and it will not be developed further. The non-commercial availability of the LDDB,[5] with a good data handling rationale, and an excellent, annually updated reference file, fills this need very satisfactorily.

However, this system could be improved by further developing its list of signals, which is the language connecting the consultation patient with the reference file. Suggested changes include the following. (1) Eliminating the maximum number of 28 signals per syndrome because, the signals used being very discriminating, the reference file is unable to include the whole expressivity of some syndromes. (2) Expanding the list of signals, including not only different levels as

it does now, but also the possibility of 'aggregated'[8] signals for well defined malformation complexes (sequences, developmental field complexes, etc), such as holoprosencephaly, Klippel-Feil syndrome, or frontonasal dysostosis. (3) Avoiding the use of synonyms, at least from the paediatrician's standpoint, such as short palpebral fissures and blepharophimosis. These changes would avoid some divergences between the undiagnosed patient and the LDDB.

This work was partially financed by the national research councils of Argentina (CONICET) and Brazil (CNPq), by FINEP, and by PAHO (Pan-American Health Office).

1 Miller RA, People HE, Myers JD. INTERNIST-I: an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med* 1982;**307**:468–76.
2 Gouvernet J, Caraboeuf M, Ayme S. GENDIAG: a computer-assisted facility in medical genetics based on belief functions. *Methods Inf Med* 1985;**24**:177–80.
3 Schorderet D, Aesbischer P. SYNDROC: microcomputer-based differential diagnosis of malformation patterns. *Arch Dis Child* 1985;**60**:248–51.
4 Schorderet D. Diagnosed human malformation patterns with a microcomputer: evaluation of two different algorithms. *Am J Med Genet* 1987;**28**:337–44.
5 Winter RM, Baraitser M, Douglas JM. A computerised data base for the diagnosis of rare dysmorphic syndromes. *J Med Genet* 1984;**21**:121–3.
6 Smith DW. *Recognizable patterns of human malformation*. 3rd ed. Philadelphia: Saunders, 1982.
7 Crisci JV, Lopez-Armengol MP. *Introduccion a la Teoria y Practica de la Taxonomia Numerica*. Washington: Secretaria General de la OEA Programa Regional de Desarrollo Cientifico y Tecnologico, 1983.
8 Eddy DM, Clanton CH. The art of diagnosis. *N Engl J Med* 1982;**306**:1263–8.