# UC Irvine

## UC Irvine Previously Published Works

**Title**

A hierarchical classes analysis (HICLAS) of primary care patients with medically unexplained somatic symptoms.

**Permalink**

https://escholarship.org/uc/item/3bz8c1mq

**Journal**

Psychiatry research, 81(1)

**ISSN**

0165-1781

**Authors**

Gara, MA
Silver, RC
Escobar, JI
et al.

**Publication Date**

1998-10-01

**DOI**

10.1016/s0165-1781(98)00088-2

**Copyright Information**

Peer reviewed

# A hierarchical classes analysis (HICLAS) of primary care patients with medically unexplained somatic symptoms

Michael A. Gara[a,*], Roxanne Cohen Silver[b], Javier I. Escobar[a],
Alison Holman[b], Howard Waitzkin[c]

[a]*Department of Psychiatry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ 08854-5635, USA*
[b]*Department of Psychology and Social Behavior, School of Social Ecology, University of California, Irvine, CA, USA*
[c]*Division of Community Medicine, University of New Mexico, Albuquerque, NM, USA*

## Abstract

This study used a clustering model, Hierarchical Classes Analysis (HICLAS), to examine patient groupings in a multiethnic sample of 1456 patients using primary care services at a university-affiliated community clinic in southern California. Somatic symptoms, psychiatric diagnoses and disability were studied using a survey instrument that included portions of the Composite International Diagnostic Interview (CIDI), the Diagnostic Interview Schedule (DIS) and the RAND-MOS Short Form Health Survey's (SF-36) 'physical functioning' dimension. HICLAS identified 11 clusters of patients with distinct patterns of medically unexplained somatic symptoms. These patient clusters varied with respect to psychiatric diagnoses and symptoms, gender, immigration status and disability. Results of this study suggest that the type of presenting symptom(s) and their various combinations may have diagnostic and prognostic value in primary care settings. These new findings may lead to further refinement of current diagnostic constructs for somatizing syndromes.   © 1998 Elsevier Science Ireland Ltd. All rights reserved.

*Keywords:* Somatoform disorders; Somatization; Hypochondriasis; Psychiatric diagnosis; Clustering analysis

*Corresponding author. Tel.: +1 732 2354351; fax: +1 732 2353258; e-mail: gara@umdnj.edu

# 1. Introduction

Somatizing syndromes have been described throughout the years under protean labels (e.g. 'ennui', 'hysteria', 'hypochondriasis', 'neurasthenia') shaped by the medical model dominant at the time (Shorter, 1993). Historically, regardless of fashion or prevailing paradigms, somatic presentations have had a place of their own in descriptive psychopathology.

Because of the pejorative connotation of some old terms (e.g. 'hysteria'), the word 'somatization', denoting psychological causality, was incorporated into clinical discourse. In current diagnostic systems, 'somatoform' became the term used to signify an overarching category that subsumed somatization disorder as well as other disorders, such as hypochondriasis, that involved medically unexplained physical symptoms. However, the acceptance of the two terms remains confined to the area of psychiatry and clinical psychology. Because hypochondriasis and somatization disorder (the two most distinctive and valid somatoform diagnoses) have low prevalence rates and fail to capture the large majority of patients presenting with unexplained medical symptoms, there is a need to develop systems of classification that are more 'user-friendly' and that can be shared with and accepted by primary care physicians.

## 1.1. Lists of somatic symptoms

The traditional lists of somatic symptoms used to elicit diagnoses of hysteria, and later somatization disorder, were actually quite comprehensive. Originally, these included not only somatic symptoms, but also other attitudinal and clinical features often seen in these patients such as dramatic demeanor, depressive, anxiety, and even psychotic symptoms. For example, in the 1960s, Perley and Guze (1962) used a list with 59 different symptoms to diagnose hysteria in their prospective studies. In the more recent classifications (DSM-III, DSM-III-R and DSM-IV), the symptoms have been restricted to somatic manifestations. The total number of 'possible' symptoms has decreased from 59 to 37 in DSM-III, 35

in DSM-III-R and now 34 in DSM-IV, and the symptom cut-off for defining a case has followed suit.

Many of the somatic symptoms listed in the above nomenclatures and included in diagnostic instruments such as the Diagnostic Interview Schedule (DIS) (Robins et al., 1981) and the Composite International Diagnostic Interview (CIDI) (Robins et al., 1988) are relatively common in general and clinical populations (e.g. gastrointestinal, cardiorespiratory). Other symptoms, while rare, provide a distinctive rubric of psychopathologic 'caseness' (e.g. pseudoneurologic or 'conversion' symptoms), and therefore, remain useful for systematic studies.

## 1.2. Somatic symptom typologies

Clustering methods have been used extensively to inform taxonomic work in psychiatry in general as well as in other areas of medicine and in biological science (such as the identification of species). In addition to clinical observations, empirical clustering studies of symptoms and signs have proved useful for the purpose of developing new diagnoses and/or validating old ones. In a thorough review of clustering applications in psychiatry, Blashfield (1986) cited over 500 such studies.

To our knowledge, only two studies have studied empirically how somatic symptoms (among other psychiatric symptoms) cluster, and both studies were carried out in non-clinical populations. Both investigations used computerized clustering algorithms to examine clustering of DIS-elicited somatic symptoms, as well as psychiatric symptoms such as anxiety and depression. The first of these studies (Swartz et al., 1986) used a procedure developed for the analysis of medical classifications called 'grade of membership analysis' or GOM (Woodbury and Manton, 1982). GOM produces 'fuzzy' as opposed to the discrete categories generated in the more traditional clustering models. According to the developers of GOM, the use of fuzzy sets may be a more appropriate way to represent 'gradations' in the individual expression of symptoms, which are so characteristic of psychiatric syndromes and classi-

fications. The sample used for GOM analysis included approx. 4000 community respondents interviewed with the DIS as part of the ECA study in the Piedmont region of North Carolina. The version of the DIS used in the study asked about and probed 47 individual symptom items for eliciting a diagnosis of somatization disorder. Of these items, 38 were somatic symptoms, four were symptoms of sexual dysfunction, two were measures of 'sickliness' and three were symptoms of depression. To this list the investigators added one psychotic symptom (hallucination or delusion), three depressive symptoms and one anxiety symptom ('nervousness') in efforts to round up all possible features of somatization disorder as described in DSM-III. Hence, the final list consisted of 52 symptoms, thus approximating the original list used by the Washington University group (Perley and Guze, 1962). Respondents endorsing at least three of the symptoms (approx. 45% of the sample) were included in the clustering analyses of somatic symptoms.

The results of the GOM analyses yielded seven 'pure' types of 'naturally occurring' somatic symptom clusters. These clusters had different symptom admixtures and sociodemographic character-

istics. A brief outline of these clusters is provided in Table 1.

In the other study, Rubio-Stipec et al. (1989) analyzed DIS interviews in approx. 3500 respondents at the Los Angeles ECA and Puerto Rican survey sites. The authors used factor analysis — which has an underlying dimensional model — as a way to represent symptom clusters. These analyses yielded five clusters in the Puerto Rican sample (alcohol, affective, phobic, psychotic and somatization disorder symptoms), all of which, except somatization, could be replicated in the Los Angeles sample. The somatization cluster found in the Puerto Rican community included 14 symptoms, five pseudoneurologic (dizziness, amnesia, fainting, paralysis, double vision), four gastrointestinal (abdominal pain, vomiting, nausea, excessive gas), two cardiorespiratory (chest pain and palpitations), one musculoskeletal (muscle weakness) and one non-specific (lifetime sickliness).

## 1.3. HICLAS clustering model

The present study makes extensive use of the Hierarchical Classes Model (HICLAS) to repre-

Table 1
GOM-elicited somatic symptom clusters (modified from Swartz et al., 1986)

| | Somatic symptom type (N) | Level of unexplained symptoms | Major symptoms | Gender | Health services use, percent making two or more visits in last 6 months (%) |
|---|---|---|---|---|---|
| I | 660 | Low | Sexual dysfunction headache | 70% female | 37 (low) |
| II | 193 | Intermediate | Female reproductive | Only female | 35 (low) |
| III | 293 | Intermediate | Depressive | Only female | 45 (intermediate) |
| IV | 85 | High | Pseudoneurologic plus many other | Only male | 81 (high) |
| V | 50 | High | Pseudoneurologic plus many other | Only female | 100 (high) |
| VI | 231 | Low | Gastrointestinal | 70% female | 44 (intermediate) |
| VII | 114 | Intermediate | Cardiorespiratory | 65% male | 78 (high) |

sent how primary care patients cluster with respect to groupings of medically unexplained somatic symptoms. HICLAS is a recently developed two-way, two-mode clustering model, well suited to binary data such as in the present application (De Boeck and Rosenberg, 1988; Gara et al., 1992; Van Mechelen et al., 1995; Rosenberg et al., 1996). The model has the capability of representing clustering in the rows (in this application, patients) and columns (symptoms) of a binary data array. Most other (but not all) clustering algorithms represent either row clusters or column clusters, but not both. In addition, HICLAS is unique in that it explicitly represents set-theoretical (superset–subset) relations among the row and column clusters.

The first published application of the HICLAS model in the area of psychiatric nosology was an analysis of how symptoms were distributed among disorders in the DSM-III-R manual itself (Gara et al., 1992). In effect, this was a quantitative analysis of the underlying taxonomy tacit in the DSM-III-R. This HICLAS analysis revealed not only several well-defined discrete symptom classes (e.g. delusions, depression, somatic symptoms) in the DSM-III-R but also several well-defined clusters of psychiatric disorders. The latter were modeled in HICLAS as combinations of one or more symptom classes and seemed to match traditional categories in descriptive psychopathology that have been incorporated in psychiatric nosologies throughout the years. Listed in hierarchical order on the basis of how well they were defined by the HICLAS program, these syndromes included: Psychoses; Mood Disorders; Organic Mental Disorders; Sleep Disorders; Addictive Disorders; Somatoform Disorders; Schizophrenia; and Anxiety Disorders. Interestingly, other less 'traditional' categories, such as Adjustment Disorders, Childhood Disorders, Personality Disorders and Sexual Disorders were not well-defined and fit the HICLAS model quite poorly. An interesting validation of the HICLAS model was the finding that the 'goodness of fit' (Jaccard measure) of disorder classes to the HICLAS model was significantly correlated to their inter-rater reliability in the DSM-III-R field trials (Gara et al., 1992).

## 2. Method

### 2.1. Subjects

The subject sample consisted of 1456 new patients who sought primary care services at a University-affiliated outpatient clinic (North Orange County Community Clinic) located in Anaheim, CA. Following completion of informed consent procedures, and in temporal proximity with their clinical examination by a physician, the patients participated in a structured interview administered by trained bilingual interviewers that included detailed questions on general demographics, psychopathology, and physical functioning. Fifty percent of those patients initially approached for the study agreed to participate. There were no demographic differences between study participants and those who declined participation, except for level of education. Those who agreed to participate had, on average, one more year of education than those who did not.

### 2.2. Instruments

Assessment of psychopathology was made with the Composite International Diagnostic Interview (CIDI) (Robins et al., 1988). Diagnoses examined included: Somatization Disorder; Generalized Anxiety; Dysthymia; and Major Depression (including melancholic subtypes). In addition, the 'physical functioning' dimension of the RAND-MOS Short Form Health Survey (SF-36) was used as a measure of disability (Brook et al., 1979). Total scores in this dimension range between 10 (severe disability) and 30 (no disability).

Bilingual (Spanish/English) research interviewers were trained in the use of the CIDI, adhering to the official CIDI training guidelines as done at the US training site located in the Department of Psychiatry, Washington University in St. Louis. All instruments were translated, pre-tested and adapted for use with Spanish-speaking subjects.

Following the standard probing system in the CIDI, symptoms were scored as 'present' if they met the severity criteria and remained medically unexplained after detailed questioning. For exam-

ple, if the respondent answered 'yes' to the question 'have you ever had abdominal or belly pain?', the interviewer proceeded with a specific set of questions to determine symptom severity, which included probes regarding physician visits, medication intake, or significant interference with daily life or functioning. If these criteria were met, the interviewer asked about the physician's diagnosis and probed whether the symptom was ever due to physical illness or injury, or followed the use of medications, drugs or alcohol. If these inquiries proved negative for medical explanations, the symptom was scored as a positive somatization symptom. Obviously, the four female reproductive items were skipped in the case of male patients. Thus, there were only 37 symptoms applicable to males.

### 2.3. HICLAS and other statistical analyses used in this study

A hierarchical classes approach to psychiatric taxonomy presupposes the assemblage of a matrix of patients × symptoms. In the present analyses, the matrix has 1455 rows corresponding to the 1455 patients with complete symptom data and 40 columns representing the various CIDI somatic symptoms. A cell entry in the matrix is '1' if the patient has the symptom and if the symptom is judged to be disruptive and remains medically unexplained after medical consultation. Otherwise the cell entry is '0'. In this model various relationships among symptoms and across subjects can be represented. For example, two or more patients can be allocated by HICLAS to the same patient cluster if they share the same pattern of symptoms (e.g. symptoms from the same organ systems). Similar patterns can also be detected for the various symptoms, and these patterns can be represented as symptom clusters that are arrayed hierarchically. Thus, HICLAS defines two hierarchical class structures for a patient × symptom matrix. One represents cluster relations among patients; the other, cluster relations among symptoms. The two structures and their relations constitute a formal hierarchical classes model for the entire matrix.

Because the resulting solution from the pri-

mary HICLAS analyses yielded an extraordinary number of low frequency ($N = 1-2$) patient classes, we also performed a second order HICLAS analysis on the results of the first analysis, in order to reduce the number of low frequency clusters. This second order analysis, analogous to a second-order factor analysis, led to the merging of various low-frequency symptom clusters that had similar patterns into larger clusters, thus bypassing the need to collapse these low frequency clusters by eyeballing or by other a priori, heuristic procedures.

The results of HICLAS analyses were related statistically (e.g. using *t*-tests) to psychiatric diagnoses such as major depression, melancholic depression, and generalized anxiety, and to disability. For the latter, an index of 'disability' was derived from the RAND-MOS 'physical functioning' dimension. Scores in this scale range between 10 (no disability) and 30 (severe disability).

## 3. Results

The patients were 55% female, and their ages ranged between 18 and 67 years (mean = 36.4; S.D. = 11.8). The sample included predominantly four ethnic groups: US Non-Hispanics ($N = 533$, a majority of them white); US born Hispanics ($N = 204$, a majority of Mexican origin); Mexican immigrants ($N = 593$) and Central American immigrants ($N = 125$, most of them from El Salvador and Guatemala). One of the 1456 subjects had incomplete data and was dropped from all further analyses reported in this article. The average number of years of completed schooling was 9.9 (S.D. = 4.1). Not surprisingly, immigrants had less schooling (7.3 years) than non-immigrants (12.6).

### 3.1. First order HICLAS analyses

A partitioning of the 41 CIDI somatic symptoms into eight clusters (as described below) was used as the initial configuration for the HICLAS analyses of the $1455 \times 41$ matrix. This a priori grouping of symptoms is necessary for HICLAS analysis. That is, HICLAS requires an initial con-

figuration (clustering) of either the row or the columns of a two-way matrix in order to minimize the possibility of local minima in the final solution (De Boeck and Rosenberg, 1988). As the focus here is on patient clusters and their properties, we thought it better to make a priori assumptions about symptom clusters than about patient clusters.

We grouped the 41 symptoms into eight specific organ/body systems as follows: (a) *pseudoneurological* (15 symptoms, such as 'blindness' and 'paralysis'); (b) *gastrointestinal* (seven symptoms); (c) *musculoskeletal* (four symptoms); (d) *genitourinary* (four symptoms); (e) *female-reproductive* (four symptoms); (f) *cardiorespiratory* (three symptoms); (g) *headache and other pain* (two symptoms: 'headache', 'and pain not classified elsewhere'); and (h) *skin* (one symptom: 'blotches or discoloration').

The resulting HICLAS solution fits the original data matrix fairly well ($\kappa = 0.73$). HICLAS also calculates a Jaccard measure of fit for each individual symptom, which is interpreted in a way that is roughly comparable to interpreting a factor loading. For example, the goodness of fit of 'blurred vision' to its associated cluster is 0.385. The higher the fit, the more representative the symptom is for the particular class or cluster in which it is placed. Some organ system clusters (headache, genito-urinary, cardiorespiratory, fe-

male-reproductive, musculoskeletal) generated particularly high fits ($> 0.50$). However, pseudo-neurological symptoms had lower fits and the 'lone' skin symptom from the CIDI did not fit the HICLAS model at all. This may be a consequence of the very low overall prevalence of these symptoms, their heterogeneity, and in the case of the skin symptom, rare co-occurrence with other symptoms in the data set.

## 3.2. Second order HICLAS analyses

These analyses allocated the 1455 patients into one of 11 major clusters, effectively reducing the large number of clusters obtained in the first-order analyses and making the clustering results more manageable.

The 11 patient clusters identified by HICLAS are labeled 'A–K' in Fig. 1. Starting at the bottom of the figure, the reader will notice six clusters (boxes F, G, H, K, I, J) corresponding to patient clusters. As is characteristic of HICLAS, patient clusters are always defined with respect to symptom clusters. For example, patient Cluster F consists exclusively of patients who have only cardiorespiratory (CR) symptoms, Cluster G of patients who have only genito-urinary (GU) symptoms, and so forth. The notation '$n = 52$' within Cluster F means that 52 patients exhibited this class of symptoms. Note that Cluster 'K' is
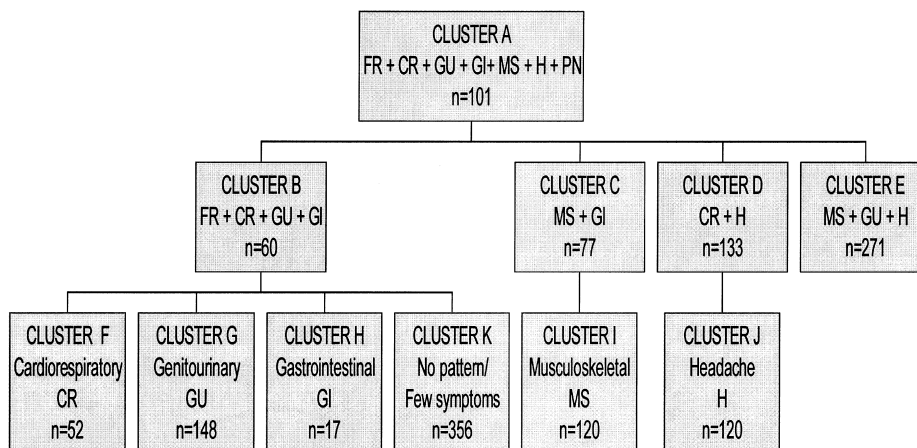


Fig. 1. Clustering of primary care patients ($N = 1455$) based on somatic complaints.

labeled as 'no pattern/few symptoms'. This represents 356 patients for whom there were either no symptoms reported, or for whom the few symptoms that were reported showed no clear co-occurrence pattern (according to HICLAS) across patients.

In addition to bottom classes (F−K), the figure also has several superordinate or 'superset' classes (A−E). For example, Cluster B represents 60 patients who have female reproductive symptoms (FR) as well as cardiorespiratory (CR), genitourinary (GU) and gastrointestinal symptoms (GI). Cluster A consists of 101 patients who have all symptom classes, including pseudoneurological symptoms (PN). The way Fig. 1 is configured also serves to illustrate that pseudoneurological symptoms tend not to occur in patients by themselves as distinct clusters, but instead, always co-occur with other symptom clusters.

### 3.3. Implications of the HICLAS model of unexplained somatic symptoms: the case of pseudoneurological symptoms

Two logical propositions (P1 and P2) can be constructed based on the HICLAS configuration in Fig. 1. The first proposition, dubbed 'P1', is the following: *when there are several pseudoneurological symptoms present, it is likely that a patient will meet Escobar et al.'s (1989) criteria for abridged somatization* (have four or more unexplained physical symptoms if male; six or more if female). This proposition P1 is based on the fact that in Fig. 1 a patient with several pseudoneurological symptoms is likely to be found in Cluster A, and patients in Cluster A also have **all** symptom clusters that are beneath Cluster A in the hierarchy. The second proposition based on the figure, dubbed P2, asserts that the relation between pseudoneurological symptoms and abridged criteria is *asymmetric*: *given that a patient meets abridged criteria, the likelihood that he or she will also have several pseudoneurological symptoms is smaller than the likelihood that a patient will meet abridged criteria given the presence of several pseudoneurological symptoms*.

In order to test whether propositions P1 and P2 were valid, and not some artifact of HICLAS analysis, we assessed the likelihoods associated with P1 and P2 using an alternative set of statistical analyses. These included cross-tabulation analysis coupled with the Somer's $D$ statistic. The latter statistic explicitly represents asymmetric predictive relationships when such are present in actual data. The cross-tabulation analysis that we used involved two binary variables: (a) a variable indicating whether or not a given patient met abridged criteria for somatization; and (b) a variable indicating whether or not a patient had three or more pseudoneurological symptoms. A total of 111 patients (7.6% of the sample) had three or more pseudoneurological symptoms; we chose three symptoms as the cutoff because patients in Cluster A averaged 2.4 such symptoms.

The results of the cross-tabulation described above were as follows. The value of Somer's $D$ was 0.74 ($P < 0.001$) in that instance when meeting abridged criteria was 'predicted' by the presence of three or more pseudoneurological symptoms. The value of $D$ was only 0.30 the other way around (i.e. 'predicting' three or more pseudoneurological symptoms when abridged criteria were met). The magnitude of the first Somer's $D$ (0.74), as well as the asymmetry of the two $D$ statistics when considered together (0.74 vs. 30), validates propositions P1 and P2. Hence, the results confirm the utility of pseudoneurological symptoms in flagging cases of somatoform disorder.

It is possible that eliciting pseudoneurological symptoms will prove to be a parsimonious way to screen for DSM-IV Somatization Disorder (SD), a diagnosis for which only eight of the 1455 patients in the present study met the criteria. The presence of three or more pseudoneurological symptoms identifies half ($n = 4$) of these SD patients, with a false positive rate of 7.4% and a false negative rate of 0.3%. An index based on four or more pseudoneurological symptoms performs even better as a screen, flagging the same four SD patients, but yielding lower rates of false positives (3.6%) and an identical rate of false negatives (0.3%). This four-symptom pseudoneurological index also predicts abridged criteria for somatization even more strongly than the three-symptom index (0.81 vs. 0.74).

We also compared, using an additional set of variables, the eight patients who met criteria for SD with the 53 patients who did not meet criteria for SD but who did report having four or more pseudoneurological symptoms. We found no statistically significant or even nominal differences between the two types of patients in terms of number of depressed symptoms, total score on the SF-36 physical functioning scale, and number of anxiety symptoms ($P$ values associated with all $t$-tests were $\geq 0.40$). Interestingly, the percentage of patients presenting four or more pseudoneurological symptoms who met criteria for lifetime major depression was 58.4%; while the percentage of patients with SD who met depression criteria was 62.0%. Again, the between-group differences in lifetime depression were not statistically significant (Fisher's exact $P = 0.83$). In fact, over 82% of the pseudoneurological patients and 87% of the SD patients met criteria for at least one lifetime DSM-IV axis I diagnosis other than a somatoform disorder. Of the 487 patients (33% of the total sample) who were diagnosed with any DSM-IV axis I diagnosis, the 53 patients who met the pseudoneurological criteria were more disabled according to the SF-36 data than those patients who did not meet the criteria [$t$ (483) = 3.37, $P < 0.001$].

### 3.4. Closer scrutiny of the patients in HICLAS Cluster A

One way of validating the HICLAS solution is to compare patients with multiple pseudoneurological and other somatic symptoms (i.e. the patients in Cluster A) with patients in the other HICLAS clusters, on a variety of axis I, symptom and demographic variables. Table 2 shows this comparison. As Table 2 indicates, Cluster A patients are quite distinct from the other primary care patients. That is, Cluster A patients are 2–3 times more likely to have a lifetime axis I diagnosis, particularly major depression. These patients also evince significantly greater physical disability, as well as a greater number of depressed and anxious symptoms. In terms of demographics, these patients are more likely to be female, born

Table 2
Comparison of Cluster A patients with others on diagnostic, symptom and demographic variables

|  | Patients in Cluster A ($N = 101$) | Other patients ($N = 1354$) |
| --- | --- | --- |
| Lifetime axis I diagnoses |  |  |
| 1. Major Depression | 44.6% | 17.0% |
| 2. Anxiety Disorder | 12.9% | 3.5% |
| 3. Dysthymic Disorder | 10.9% | 3.8% |
| 4. Melancholia | 15.8% | 4.5% |
| 5. Any of the above | 56.4% | 22.1% |
|  |  |  |
| Demographic variables |  |  |
| 6. Females | 82.2% | 52.9% |
| 7. Immigrants | 24.8% | 51.2% |
| 8. Age | 42.0 (11.1) | 36.0 (11.8) |
| 9. Years of schooling | 10.3 (3.9) | 9.9 (4.1) |
|  |  |  |
| Symptom counts; disability |  |  |
| 10. Depressed symptoms | 4.0 (3.1) | 1.7 (2.5) |
| 11. Anxious symptoms | 3.5 (5.9) | 0.9 (3.0) |
| 12. Physical functioning | 20.3 (5.6) | 25.0 (5.7) |

*Notes.* Significance levels for all variables were $< 0.0001$ except for Years of schooling ($P < 0.31$) and Dysthymic Disorder ($P < 0.001$). The $t$-tests were used to test significance for variables 8–12; Fisher's exact test was used for variables 1–7. Standard deviations are in parentheses, where applicable.

in the US, and somewhat older than the patients in other clusters.

### 4. Discussion

This study represents the first application of HICLAS for examining how primary care patients cluster with respect to medically unexplained somatic complaints. Eleven patient clusters were identified with distinct patterns of somatic symptoms. The HICLAS model that identified these clusters also fit the data fairly well. In addition, it was demonstrated in this study that an asymmetric relationship held between numerous pseudoneurological symptoms and somatization (abridged concept), such that the former predicted the latter but not vice-versa. We were able to detect this asymmetry because patients with several pseudoneurological symptoms were lo-

cated by HICLAS in a superset class (i.e. Class A) at the upper-most level of the hierarchy, subsuming all clusters at lower levels. This important role of pseudoneurological symptoms was also confirmed by analyses showing that the presence of four or more such symptoms was strongly associated with somatization, whether defined by DSM-IV criteria or by abridged criteria.

With respect to the external validity of the HICLAS result, it was found that patients with pseudoneurological and multiple other somatic symptoms (Cluster A) differed from other patients in terms of psychiatric comorbidity, physical disability, and demographic factors. This clearly represents a more severe form of somatization, less likely to be seen in the case of immigrant patients compared to the US born. Naturally there are certain characteristics of the study (e.g. attrition, unique ethnic makeup of primary care patients sampled, and use of lay interviewers) that limit the generalizability of the clustering results to all populations.

### 4.1. The relationship of HICLAS to grade of membership (GOM) analysis

It is interesting to note that the current results using HICLAS replicate, at least in part, those of Swartz et al. (1986) using a different analytic technique (GOM) on a general population sample. For example, their 'pure' type V (see Table 1) very closely resembles our HICLAS Cluster 'A' (Fig. 1). Thus, both of these clusters are seen almost exclusively among females (Cluster A is 82% female), include high levels of unexplained symptoms coming from multiple organ systems and augur high levels of disability or health services use. Also, their 'pure' type I may be akin to HICLAS Cluster 'K' ('no pattern/few symptoms'), and their single organ symptom clusters (types VI and VII) seem very similar to HICLAS Clusters 'F' and 'H', which were made up almost exclusively of gastrointestinal and cardiorespiratory symptoms.

HICLAS is based on an underlying discrete model of category membership, as opposed to the GOM model, which is based on fuzzy set theory. The novelty of HICLAS lies in the fact that hierarchical (superset/subset) arrangements among classes are explicitly represented. For example, we have seen already that the location of pseudoneurological symptoms in a superordinate class ('A') by HICLAS has important implications for classifying somatoform syndromes. However, while GOM does locate pseudoneurological symptoms in two distinct classes, a male class (IV) and a female class (V), and does show their high levels of co-occurrence with other somatic symptoms, the representation of superset-subset relationships between pseudoneourological and other symptom clusters is beyond the purview of GOM analysis. Nonetheless, GOM analysis is quite useful in other respects, and the fact that there is considerable convergence between the symptom clusters (e.g. pseudoneurological; cardiorespiratory) identified by GOM in one large general population sample and those identified by HICLAS in another large, albeit clinical, sample bodes well for the development of a method-independent, general typology of medically unexplained somatic symptoms.

### 4.2. Revised somatization construct

The present data may contribute to the further refinement of diagnostic indices of somatization such as the abridged somatization construct (Escobar et al., 1989) by adding type and number of organ systems to the more generic high levels of unexplained symptoms. These new observations may provide the construct more precision in discriminating among various psychopathologies, thus improving the detection of 'pure' and 'mixed' somatizing syndromes. In addition, the finding that four or more pseudoneurological symptoms was highly predictive of DSM-IV criteria for SD and abridged criteria for somatoform disorder is intriguing, and deserves replication in other large samples of primary care patients. While awaiting such future research, a tentative suggestion for primary care providers is to pay close attention to medically unexplained pseudoneurological symptoms. Upon observing the co-occurrence of several of these symptoms in a patient, the provider should consider a somatoform diagnosis as well other co-morbid axis I diagnoses, such as major

depression. He or she should also be apprised of the fact that the coupling of a psychiatric diagnosis with several pseudoneurological symptoms is associated with considerable functional impairment.

## References

Blashfield, R.K., 1986. Structural approaches to classification. In: Million, T., Klerman, G.I. (Eds.), Contemporary Directions in Psychopathology: Toward the DSM IV. The Guilford Press, New York, pp. 363–380.

Brook, R.H., Ware, J.E., Stewart, A.L., Donald, C.A., Rogers, W.A., Williams, K.N., Johnson, S.A., 1979. Conceptualization and Measurement of Health for Adults in the Health Insurance Study, vol. VII, Overview. Rand Corporation, Santa Monica, CA.

De Boeck, P., Rosenberg, S., 1988. Hierarchical classes: model and data analysis. Psychometrika 53, 361–381.

Escobar, J.I., Manu, P., Lane, T., Matthews, D., Swartz, M., 1989. Medically unexplained physical symptoms, somatization disorder and abridged somatization: studies with the DIS. Psychiatric Developments 3, 235–245.

Gara, M.A., Rosenberg, S., Goldberg, L., 1992. DSM-III-R as a taxonomy: a cluster analysis of diagnoses and symptoms. Journal of Nervous and Mental Disease 180, 11–17.

Perley, M.J., Guze, S.B., 1962. Hysteria — the stability and usefulness of clinical criteria. New England Journal of Medicine 266, 4211–4226.

Robins, L.N., Helzer, J.E., Croughan, J., Ratcliff, K.S., 1981. National Institute of Mental Health Diagnostic Interview Schedule. Archives of General Psychiatry 35, 73–782.

Robins, L.N., Wing, J., Wittchen, H.U., Babor, T., Burke, J., Farmer, A., Jablenski, A., Pickens, R., Regier, D.A., Sartorius, N., Towle, L.H., 1988. The Composite International Diagnostic Interview. Archives of General Psychiatry 45, 1069–1077.

Rosenberg, S., Van Mechelen, I., De Boeck, P., 1996. A hierarchical classes model: theory and method with applications in psychology and psychology. In: Arabie, P., Hubert, L., De Soete, G. (Eds.), Clustering and Classification. World Scientific, Teaneck, NJ, pp. 123–155.

Rubio-Stipec, M., Shrout, P.E., Bird, H., Canino, G., Bravo, M., 1989. Symptom scales of the Diagnostic Interview Schedule: factor analytic results in Hispanic and Anglo samples. Psychological Assessment 1, 30–34.

Shorter, E., 1993. From Paralysis to Fatigue. The Free Press, New York.

Swartz, M., Blazer, D., Woodbury, M., George, L., Landerman, R., 1986. Somatization disorder in a US southern community: use of a new procedure for analysis of medical classification. Psychological Medicine 16, 595–609.

Van Mechelen, I., De Boeck, P., Rosenberg, S., 1995. The conjunctive model of hierarchical classes. Psychometrika 60, 505–521.

Woodbury, M.A., Manton, K.G., 1982. A new procedure for the analysis of medical classification. Methods of Information in Medicine 21, 210–220.