Classification of Dyspepsia

Identification of Independent Symptom Components in 7270 Consecutive, Unselected Dyspepsia Patients from General Practice

V. MEINECHE-SCHMIDT & E. CHRISTENSEN
Dept. of General Practice, The Panum Institute, and Clinic of Internal Medicine I, Bispebjerg University Hospital, Copenhagen, Denmark

Meineche-Schmidt V, Christensen E. Classification of dyspepsia. Identification of independent symptom components in 7270 consecutive, unselected dyspepsia patients from general practice. Scand J Gastroenterol 1998;33:1262–1272.

Background: Several attempts to classify dyspepsia into subgroups have been proposed as a basis for empirical treatment and research. However, subgrouping has proved difficult due to overlap of symptoms between subgroups, and the response to empirical therapy is difficult to predict. We aimed to study whether natural symptom combinations occur in patients seeing general practitioners because of dyspepsia and whether symptom presentation could predict the effect of proton pump inhibitor treatment. *Methods:* The symptom presentation of 7270 consecutive, unselected patients with dyspepsia in general practice was studied by using principal-components analysis. The relation to the effect of omeprazole was studied in a subsample (n= 471) with predominantly reflux-like or ulcer-like dyspepsia being included in a controlled clinical trial of omeprazole versus placebo. *Results:* Four principal components (factors), explaining 36% of the total variance, were found. They describe four independent dimensions in the symptoms of dyspepsia that can be interpreted meaningfully as representing A) acid-related disease of the upper gastrointestinal tract, B) irritable bowel disorder, C) dysmotility of the stomach/duodenum, and D) dysmotility of the esophagus. In the subsample the response to proton pump inhibition therapy was associated with high component-A scores, low component-B scores, and low component-C scores. A pocket chart was devised to obtain the component scores easily in new patients. *Conclusion:* The analysis identified four characteristic, biologically meaningful dyspepsia components that express independent dimensions in the symptoms of patients with dyspepsia. The symptom scores corresponding to the four components may improve symptom-based diagnosis and thereby empirical therapy. In particular, the association between component scores and the effect of omeprazole suggests that classifying dyspepsia on the basis of these components may focus empirical omeprazole therapy even more.

Key words: Dyspepsia; general practice; principal-components analysis; proton pump inhibitor; symptom clusters

Dr. V. Meineche-Schmidt, Strandvejen 554, DK-3120 Dronningmølle, Denmark (fax: +45 4971 7911)

The definition of dyspepsia has changed over time, and at least a dozen definitions have been proposed during the past 25 years (1). General population studies have shown dyspepsia to be very common, being experienced in Sweden in 54% within a 3-month period (2), in the United Kingdom in 41% within 6 months (3), and in Denmark in 50% within 1 year (4). Although the definitions used are not fully comparable, dyspepsia must be considered a very commen event. In primary health care in Denmark about 5% of all consultations relate to dyspepsia (5). In the United Kingdom the figure is 4% (6), in Holland 3% (7), and in Norway 6% (8). The substantial impact on primary health care and the fact that most patients are being evaluated and treated by general practitioners, only a few being referred to specialists (9), emphasize the relevance of improving the symptom-based dyspepsia classification to assist empirical selection of the most effective treatment. On the basis of clinical experience several working parties comprising specialists in the field (1, 10, 11) have proposed different classifications of dyspepsia on the basis of symptoms and likely underlying pathophysiology. Application of these classifications in clinical practice has been disappointing. General population studies on the basis of questionnaires have shown extensive overlap between subgroups (12, 13), making a classification difficult. In patients referred for endoscopy no correlation between dyspepsia subgroups and endoscopic findings has been found (14). Thus, the value of a priori defined dyspepsia subgroups has not been shown.

Classification of dyspepsia has been approached in another manner (15, 16); the symptoms in a random population sample were investigated for natural symptom clustering, and three symtom clusters were found, corresponding to the irritable bowel syndrome, the upper dyspepsia–heartburn type, and the upper dyspepsia–vomiting type. No studies on the value of natural symptom clusters in predicting the outcome of empirical therapy have been reported.

Table I. Diagnostic chart. All symptoms in a patient are marked, and the dyspepsia subtype is determined by the category having the highest number of crosses

	Acid-related		Non-acid-related		
	Reflux	Ulcer	Dysmo.	Unchar.	'Alarm'
Pain where?					
Heartburn	xx				
Epigastric	x	x			
Upper abdominal			x		
Other abdominal			x	X	
Pain, when?					
In the morning				X	
At nighttime	x	x			
After meals	x		x	X	
Pain, relieved by?					
Food		xx			
Antacids	x	x			
Vomiting		x			
Stools or flatus			xx		
Other complaints					
Acid regurgitation	x				
Dysphagia	x				x
Nausea				X	
Morning vomiting				X	
Bloating			x		
Constipation			x		
Loose stools			x		
Incomplete evacuation			x		
Blood in stools					x
Black stools		x			x
Weight loss					x
Jaundice					x
Anemia					x
Total score					

The aims of this study were to study whether characteristic independent symptom components could be identified in a large sample of patients consulting a general practitioner because of dyspepsia, and to test the association of the identified symptom components with the effect of proton pump inhibition therapy in a subset of patients included in a controlled clinical trial.

PATIENTS AND METHODS

The study population comprised 7270 consecutive, unselected, uninvestigated patients presenting with dyspepsia in general practice during the period June 1991 to May 1993, as reported in detail elsewhere (17). Only patients with a complete set of variables were included in this study. Thus 4 of the 7274 patients in the original study (17) were excluded, 1 because of missing information about sex and 3 because of missing information about age. The patients had consulted a general practitioner because of dyspepsia and were subjected to a structured interview by the physician, asking the patient about the presence or absense of the symptoms listed in Table I.

To identify characteristic symptom components, the data structure of the patients was studied by using a principal-components analysis (18). This method can combine groups of correlated variables into a smaller number of independent (uncorrelated) new variables (that is, principal-components scores) explaining a large part of the variation (18). The statistical analysis is described in detail in the Appendix.

Patients with predominantly ulcer-like or reflux-like dyspepsia (having most of their symptoms in those categories (Table I)) were included in randomized clinical trials studying the effect of acid-suppressive treatment for 2 weeks, as reported in detail elsewhere (19). The relation of the principal-component scores to the effect of omeprazole was assessed in a subsample comprising 471 per-protocol-treated patients included in a randomized clinical trial comparing 20 mg omeprazole daily (243 patients) with placebo (228 patients) with complete 2-week follow-up. (Six patients were not included in this study: two omeprazole-treated and four placebo-treated patients in whom overall but not specific symptoms were recorded at the end of the treatment.)

The association between the effect of omeprazole and the principal-component scores (that is, the new independent (uncorrelated) variables produced by the analysis) was analyzed by studying the therapeutic gain (omeprazole response–placebo response) with 95% confidence intervals

Table II. Characteristics of the studied 7270 patients with dyspepsia in general practice and of the subsample of 471 patients with predominantly reflux-like or ulcer-like dyspepsia

Variable	Mean or percentage	
	Dyspepsia, $n = 7270$	Reflux-like or ulcer-like dyspepsia, $n = 471$
Mean age (years)	47.1 (s*, 18.7)	40.9 (s, 12.6)
Females	56.4%	51.2%
Epigastric pain	63.1%	89.6%
Bloating	42.0%	33.8%
Acid regurgitation	41.7%	68.6%
Pain at nighttime	41.3%	57.7%
Pain relieved by antacids	37.2%	63.5%
Pain after meals	36.2%	38.2%
Upper abdominal pain	35.8%	14.9%
Nausea	31.7%	32.5%
Pain relieved by food	29.1%	50.7%
Heartburn	26.2%	46.3%
Other abdominal pain	25.4%	2.8%
Pain relieved by stools or flatus	25.2%	9.3%
Loose stools	23.4%	16.1%
Pain in the morning	16.6%	17.4%
Constipation	16.6%	9.6%
Pain relieved by vomiting	13.2%	16.1%
Incomplete evacuation	8.3%	4.2%
Morning vomiting	4.7%	2.1%
Weight loss	3.9%	0.0%
Dysphagia	3.4%	0.0%

* s = standard deviation.

in subgroups of patients defined by one or more of the principal-component scores.

The characteristics of the patients included in the total sample and the subsample are shown in Table II. Epigastric pain was the most frequent symptom, followed by bloating, acid regurgitation, pain in the night, and pain relieved by antacids. The two samples are not uniform. Owing to the selection criteria used, the subsample had a somewhat higher proportion of acid-related dyspepsia and a lower proportion of non-acid-related dyspepsia than the total sample.

RESULTS

The four principal components obtained by analysis of the total sample are shown in Table III.

Component A was positively correlated with epigastric pain, pain relieved by antacids, acid regurgitation, heartburn, pain in the night, and pain relieved by food and negatively correlated with other abdominal pain. This component can be interpreted as representing acid-related disease involving the upper gastrointestinal tract.

Component B was positively correlated with bloating, pain relieved by stools or flatus, constipation, incomplete rectal evacuation, upper abdominal pain, and loose stools. This component can be interpreted as representing the irritable bowel syndrome.

Component C was positively correlated with nausea, morning vomiting, pain relieved by vomiting, and pain in

Table III. Principal-components factor loadings (correlations of original variables with the factors)

Variable	Component (factor)			
	A Acid-related disorder in the upper GI tract	B Irritable bowel disorder	C Dysmotility stomach/duodenum	D Dysmotility esophagus
Epigastric pain	**0.71**	−0.19	0.08	−0.04
Pain relieved by antacids	**0.68**	−0.12	−0.15	−0.14
Acid regurgitation	**0.63**	−0.10	−0.03	0.08
Other abdominal pain	**−0.63**	0.16	0.00	−0.01
Heartburn	**0.53**	−0.19	−0.17	0.24
Pain in the night	**0.52**	−0.05	0.12	−0.10
Pain relieved by food	0.49	−0.02	0.06	**−0.52**
Bloating	−0.14	**0.68**	−0.02	−0.01
Pain relieved by stools or flatus	−0.27	**0.64**	−0.05	−0.04
Constipation	−0.01	**0.55**	−0.08	0.06
Incomplete evacuation	0.02	**0.53**	0.09	−0.01
Horizontal upper abdominal pain	−0.38	0.47	0.05	0.04
Loose stools	−0.22	0.38	0.20	−0.09
Female gender	−0.08	0.20	0.08	0.14
Nausea	−0.04	0.19	**0.62**	0.16
Morning vomiting	−0.02	−0.06	**0.62**	0.05
Pain relieved by vomiting	0.23	−0.05	**0.54**	0.07
Pain in the morning	−0.05	0.14	0.38	−0.07
Age	0.29	0.15	−0.35	0.12
Pain after meals	−0.02	0.21	−0.02	**0.64**
Dysphagia	0.09	−0.10	−0.03	**0.55**
Weight loss	0.01	−0.03	0.23	0.39
Eigenvalue	3.00	2.12	1.52	1.33
Variance explained (%)	13.6	9.6	6.9	6.0

Correlations >0.5 are in heavy type in the table.

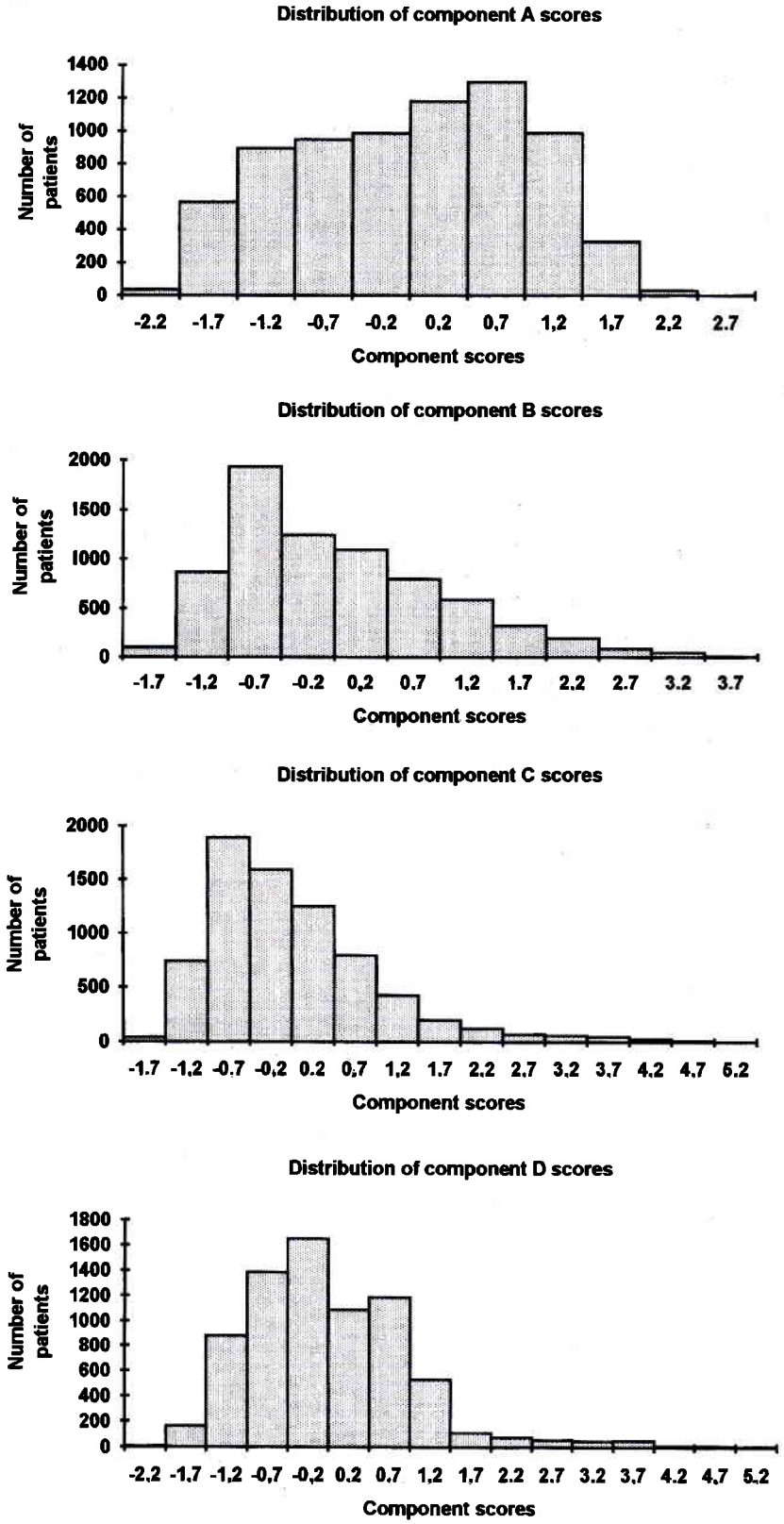

Fig. 1. Distribution of the patient scores for the four identified principal components (A–D) in the total sample of 7270 patients with dyspepsia.

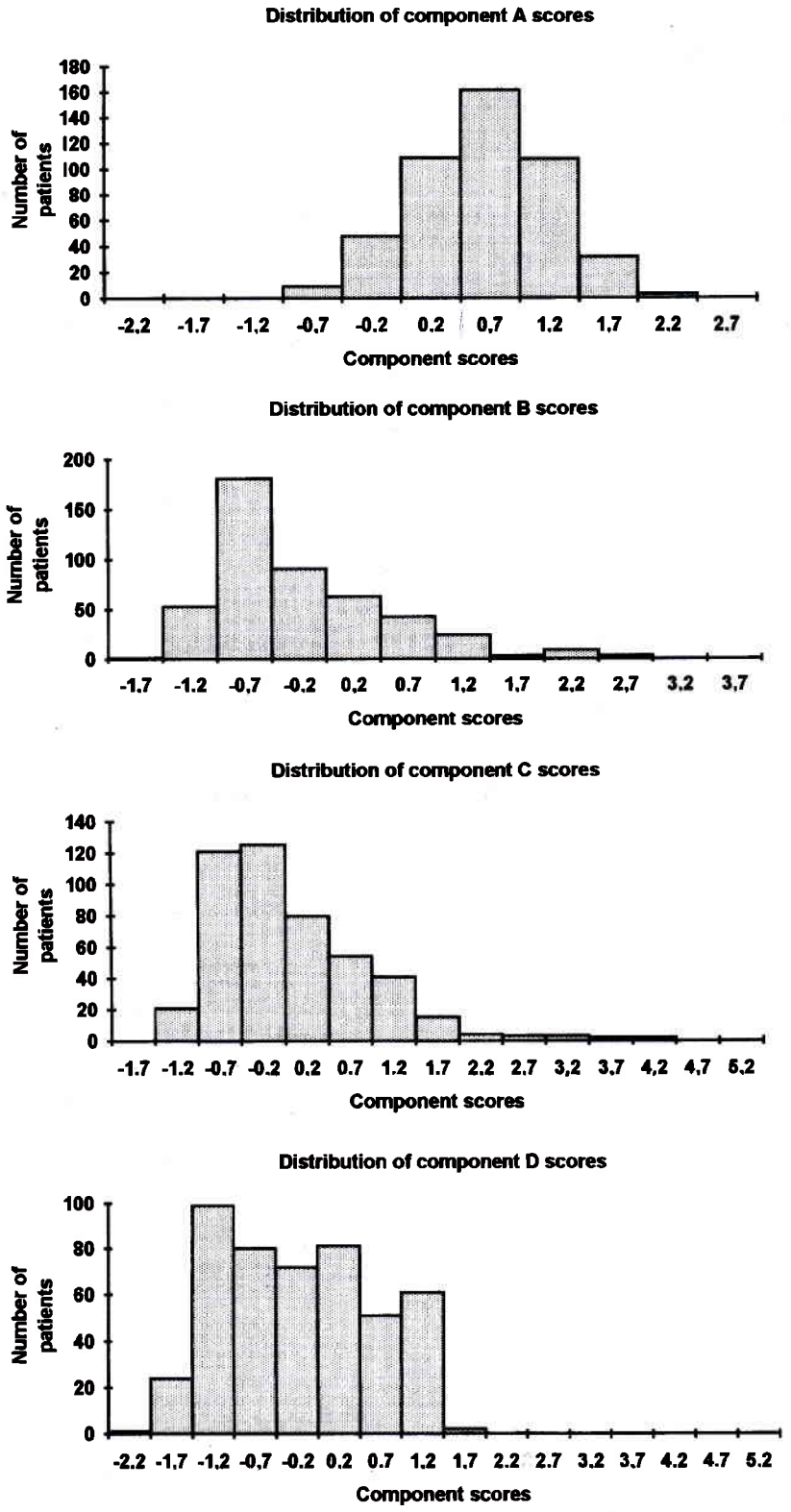

Fig. 2. Distribution of the patient scores for the four identified principal components (A–D) in the subsample of 471 patients with predominantly ulcer-like or reflux-like dyspepsia.

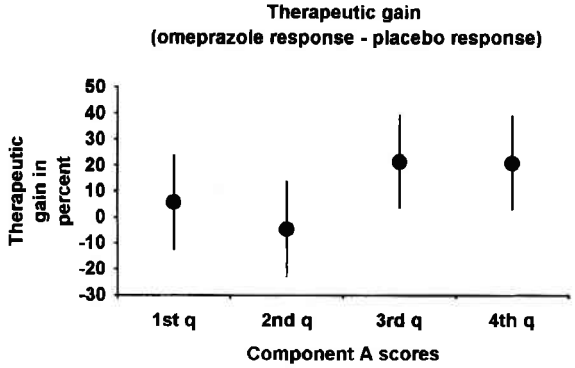

Fig. 3. The observed therapeutic gain (percentage response with omeprazole minus percentage response with placebo) with 95% confidence limits as a function of component-A scores in the subsample divided into four groups of equal size (1st to 4th quarter) on the basis of the component-A scores.

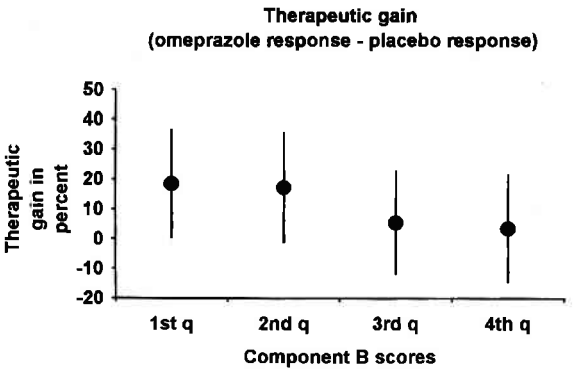

Fig. 4. The observed therapeutic gain (percentage response with omeprazole minus percentage response with placebo) with 95% confidence limits as a function of component-B scores in the subsample divided into four groups of equal size (1st to 4th quarter) on the basis of the component-B scores.

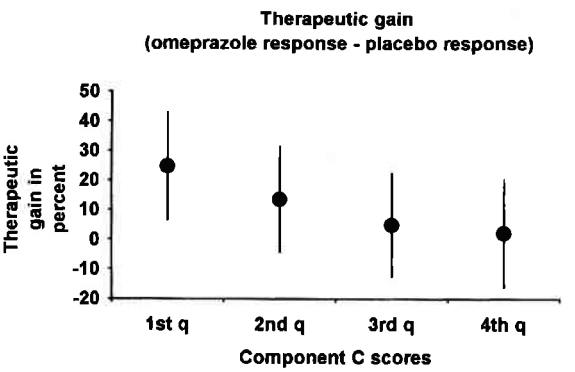

Fig. 5. The observed therapeutic gain (percentage response with omeprazole minus percentage response with placebo) with 95% confidence limits as a function of component-C scores in the subsample divided into four groups of equal size (1st to 4th quarter) on the basis of the component-C scores.

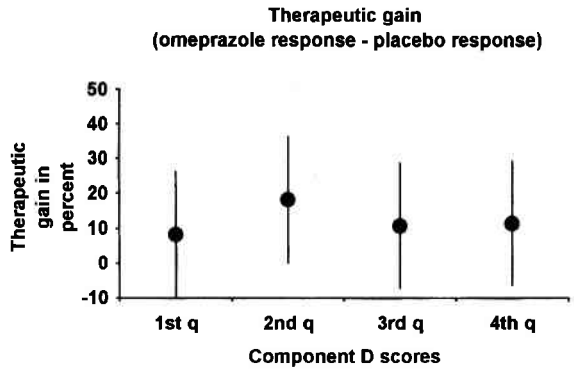

Fig. 6. The observed therapeutic gain (percentage response with omeprazole minus percentage response with placebo) with 95% confidence limits as a function of component-D scores in the subsample divided into four groups of equal size (1st to 4th quarter) on the basis of the component-D scores.

the morning. This component can be interpreted as representing dysmotility of the stomach and duodenum.

Component D was positively correlated with pain after meals, dysphagia, and weight loss and negatively correlated with pain relieved by food. This component can be interpreted as representing dysmotility of the esophagus.

The distribution of the individual factor scores for the 4 principal components are shown in Fig. 1 for the 7270 patients. No clearly delimited symptom clusters are apparent in the one-dimensional plots in the figure or in other two- or three-dimensional plots (not shown). These smooth distributions suggest that a) a given type of dyspepsia presents itself with markedly varying symptoms or symptom combinations leading to a spectrum of variations in the presentation of the patients and/or b) more types of dyspepsia may coexist in the same patient.

Fig. 2 shows the distributions of the individual factor scores in the subsample. As a consequence of the selection of the subsample the distributions are not uniform with those of the total sample, the most marked differences being that the low scores for component A and the very high scores for component D are missing in the subsample.

Figs. 3–6 show the association between the principal-component scores (A to D), divided into four groups in accordance with the quartiles of the distributions, and the therapeutic gain of omeprazole. Significant omeprazole effect appears to be associated with high component-A scores, low component-B scores, and low component-C scores. Component D showed little association with the omeprazole effect.

Figs. 7–9 show the association between pairwise principal-component score combinations and the therapeutic gain of omeprazole. A particularly high omeprazole effect (>30%) was found for high component-A score combined with low component-B score (Fig. 7) and for low component-B score combined with low component-C score (Fig. 9).

Fig. 10 shows the association between combinations of scores for components A, B, and C and the therapeutic gain of

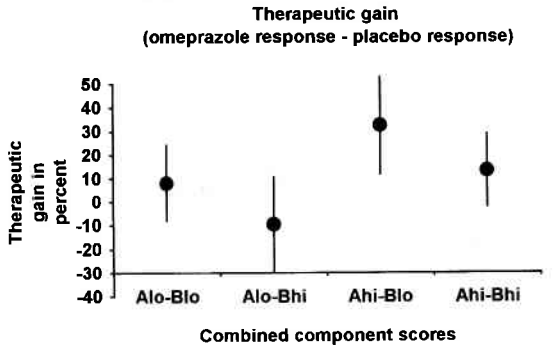

Fig. 7. The observed therapeutic gain (percentage response with omeprazole minus percentage response with placebo) with 95% confidence limits for the low (lo) and high (hi) half of component-A scores, respectively, combined with the low (lo) and high (hi) half of component-B scores, respectively.

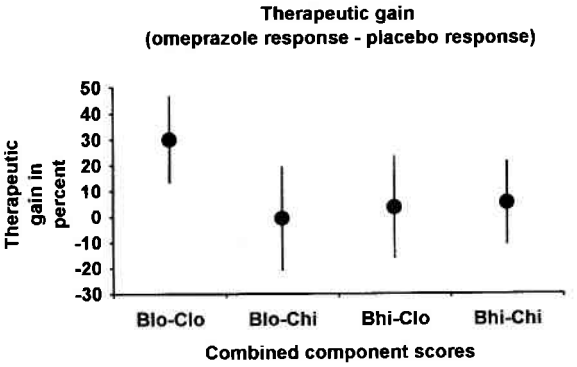

Fig. 9. The observed therapeutic gain (percentage response with omeprazole minus percentage response with placebo) with 95% confidence limits for the low (lo) and high (hi) half of component-B scores, respectively, combined with the low (lo) and high (hi) half of component-C scores, respectively.

omeprazole. The highest therapeutic gain (>40%) was found for a high component-A score combined with a low component-B score and a low component-C score.

Calculation of dyspepsia factor scores in a new patient

To facilitate application of the results in new patients, we have devised a dyspepsia chart (Table IV) by which the factor score (times 2) for each of the four dyspepsia components may be easily obtained. The details of how the dyspepsia chart was developed is described in the Appendix.

In a given patient the numbers corresponding to his/her symptoms should be added together for each of the four components. A positive sum for a factor suggests that the corresponding type of dyspepsia is probably present in the patient. If several sums are positive, then the highest sum and its corresponding type of dyspepsia should be considered the primary problem of the patient.

Examples

Example 1: A 20-year-old man presented with the following symptoms: pain relieved by antacids, acid regurgitation, heartburn, pain in the night, and pain relieved by food. Using Table 4, the score obtained for A is $1 + 1 + 1 + 1 + 1 - 2$ (the correction) = 3. The score for B is −3 (the correction). The C score is 1 (for age) − 1 (the correction) = 0. The D score is 1 (for heartburn) − 2 (for pain relieved by food) − 1 (the correction) = −3. For this rather characteristic patient the likely type of dyspepsia is A, or acid-related disease. The patient was treated with omeprazole, and the symptoms disappeared.

Example 2: A 46-year-old man presented with the following symptoms: epigastric pain, acid regurgitation, pain in the night, bloating, constipation, and pain after meals. On the basis of Table IV the score for A is 1 (for epigastric pain) + 1 (for acid regurgitation) + 1 (for pain in the night) + 1 (for constipation) − 2 (the correction) = 2. The B score is 1 (for bloating) + 2 (for constipation) − 3 (the correction) = 0. The C score is = −1 (the correction). The D score is 2 (for pain after meals) − 1 (the correction) = 1. The scores for A and D are positive, but A is highest, suggesting that acid-related disease is the predominant problem. This patient was treated with omeprazole and responded satisfactorily.

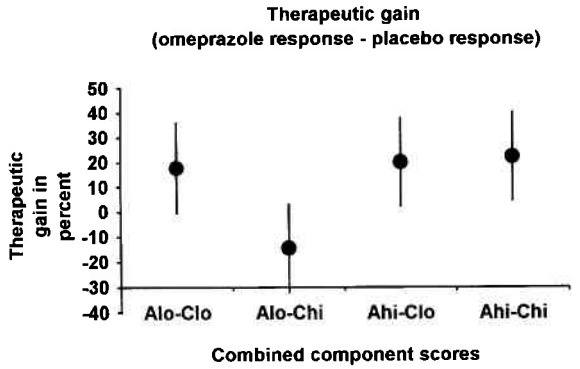

Fig. 8. The observed therapeutic gain (percentage response with omeprazole minus percentage response with placebo) with 95% confidence limits for the low (lo) and high (hi) half of component-A scores, respectively, combined with the low (lo) and high (hi) half of component-C scores, respectively.

DISCUSSION

Subgrouping of dyspepsia

Classifying patients with dyspepsia into a priori defined subgroups on the basis of working parties will reflect the clinical experience of the participating physicians, leading to differences in definitions (1, 10, 11). When applied to patients in substantially different settings, these classifications have been disappointing. It has even been suggested (20) that taking the history of the patient is a waste of time and should be avoided for the benefit of obtaining an endoscopic diagnosis in all patients with dyspeptic complaints. The vast number patients with dyspepsia (2–6) and the fluctuating

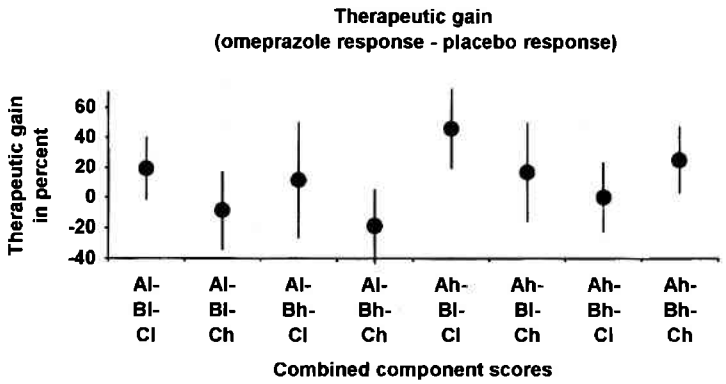

Fig. 10. The observed therapeutic gain (percentage response with omeprazole minus percentage response with placebo) with 95% confidence limits for the low (lo) and high (hi) half of component-A scores, respectively, combined with the low (lo) and high (hi) half of component-B scores, respectively, combined with the low (lo) and high (hi) half of component-C scores, respectively.

nature of the symptoms (13) makes this approach impossible, even if it is regarded as the best strategy. Most of the patients are being treated in general practice, and therefore the symptoms presented by the patients must form the basis for therapy. The disappointing results obtained by using a priori classifications calls for an improved classification of dyspepsia.

The symptoms of dyspepsia may vary considerably between patients, and individual symptom profiles may be difficult to interpret. Principal-components analysis is an explorative analytic tool that can clarify the structure of complicated data by identifying the most important independent dimensions. The analysis showed that a characteristic pattern of correlations exists between the variables and that

Table IV. Score chart for easy approximate calculation of dyspepsia factor scores ×2 according to each of the 4 principal components

Variable	Component A Acid-related disorder in the upper GI tract	B Irritable bowel disorder	C Dysmotility stomach/duodenum	D Dysmotility esophagus
Epigastric pain	1	0	0	0
Pain relieved by antacids	1	0	0	0
Acid regurgitation	1	0	0	0
Other abdominal pain	−1	0	0	0
Heartburn	1	0	0	1
Pain in the night	1	0	0	0
Pain relieved by food	1	0	0	−2
Bloating	0	1	0	0
Pain relieved by stools or flatus	0	1	0	0
Constipation	1	2	0	0
Incomplete evacuation	1	2	0	0
Horizontol upper abdominal pain	0	1	0	0
Loose stools	0	1	0	0
Nausea	0	0	2	0
Morning vomiting	0	−1	4	0
Pain relieved by vomiting	1	0	2	0
Pain in the morning	0	0	1	0
Age				
≤24 years	0	0	1	0
25–64 years	0	0	0	0
65–74 years	0	0	−1	0
≥75 years	1	1	−1	0
Pain after meals	0	0	0	2
Dysphagia	0	−1	0	5
Weight loss	0	0	1	3
Correction	−2	−3	−1	−1
Sum of points:				

* Sex had no influence on the chart calculations.

these could be combined into four new independent variables (principal-components or factors), which could be interpreted meaningfully as representing A) acid-related disease of the upper gastrointestinal tract, B) irritable bowel disorder, C) dysmotility of the stomach/duodenum, and D) dysmotility of the esophagus. However, only a small part (36.1%) of the variation in the data is explained by the principal components identified. Furthermore, classification of the patients in accordance with one or more of the components did not show distinctive, well-defined clusters. This suggests that a given type of dyspepsia presents with markedly varying symptoms and symptom combinations giving rise to a spectrum of variations in the presentation of the patients and/or that more types of dyspepsia may coexist in the same patient.

Kay & Jørgensen (15) studied dyspeptic symptoms in a population sample of 4807 men and women aged 30, 40, 50, and 60 years in 2 studies with an interval of 5 years. They recorded symptoms comparable to those in our study: abdominal pain, distention, borborygmia, altered stool consistency, heartburn, acid regurgitation, nausea, and vomiting. The symptom presentation in patients was analyzed to detect natural clustering of dyspeptic symptoms as defined by symptom combinations that occurred more often than could be explained by chance. The result was identification of three syndromes: 'irritable bowel' (abdominal pain and distention with borborygmi and/or altered stool consistency), 'upper dyspepsia, heartburn type' (abdominal pain, heartburn, and regurgitation), and 'upper dyspepsia, nausea type' (abdominal pain and nausea).

The three first components in our study correspond to the syndromes in the study by Kay & Jørgensen (15, 16) as follows: component A corresponds to upper dyspepsia, heartburn type; component B to irritable bowel syndrome; and component C to upper dyspepsia, nausea type. Our component D (dysmotility of the esophagus) was not identified as a distinctive group by Kay & Jørgensen (15, 16).

Prediction of therapeutic response

A lack of correlation between symptoms and endoscopic findings is well known; this reflects the diagnosis 'non-ulcer dyspepsia', comprising most of the patients referred to endoscopy from primary health care (14). Furthermore, it has been shown that endoscopically confirmed ulcers or esophagitis can be present without any dyspeptic symptoms (21). This may explain the lack of correlation between a priori dyspepsia subgroups and the result of empirical treatment (14).

In this study the patients presented dyspeptic symptoms severe enough to consult a general practitioner. Since only one in five patients with dyspepsia (2) will consult a GP, the results cannot be compared with population-based studies, nor can they be compared with patients referred to endoscopy, since these comprise less than 10% of the patients consulting in general practice (9).

The possibility remains that, in patients presenting in general practice because of dyspepsia, the symptoms could form a basis for therapy, making an intervention on the primary care level—that is, without a morphologic diagnosis—meaningful and worthwhile.

The association of the omeprazole effect with high component-A scores, low component-B scores, and low component-C scores supports the relevance of the identified principal dyspepsia components for the empirical decision about therapy.

Conclusion

Although the identified symptom components are not sufficiently strong to provide clearly defined clusters of patients, the study has identified four biologically meaningful, independent symptom components, which can be interpreted as representing: acid-related disease involving the upper gastrointestinal tract (component A), irritable bowel syndrome (component B), dysmotility of the stomach and duodenum (component C), and dysmotility of the esophagus (component D). The identified four independent symptom components substantiate current views about the mechanisms of the commonest causes of dyspepsia. The associations shown between the omeprazole effect and the principal-component scores A, B, and C support the relevance of the identified components. We suggest that the dyspepsia score chart based on the analysis should be used for symptom-based evaluation of dyspepsia in new patients and that its value in guiding empirical therapy should be evaluated further in new controlled clinical trials testing other therapies.

ACKNOWLEDGEMENTS

The study was supported by grants from the Public Health Insurance in Denmark, the General Practitioners Research Foundation, and the ASTRA Research Foundation.

REFERENCES

1. Talley NJ, Colin-Jones D, Koch KL, Koch M, Nyren O, Stanghellini V. Functional dyspepsia: a classification with guidelines for diagnosis and management. Gastroenterol Int 1991;4:145–60.
2. Jones RH, Lydeard SE, Hobbs FDR, Kenkre JE, Williams EI, Jones SJ, et al. Dyspepsia in England and Scotland. Gut 1990;31:401–5.
3. Agréus L, Svärdsudd K, Nyrén O, Tibblin G. Prevalence of dyspepsia and demographic characteristies in a Swedish adult population. Scand J Gastroenterol 1994;29:102–9.
4. Kay L, Jørgensen T. Epidemiology of upper dyspepsia in a random population. Prevalence, incidence, natural history, and risk factors. Scand J Gastroenterol 1994;29:1–6.
5. Wiingaard P, Ovesen L, Juul S Sygdomsmønstret i almen praksis. Aarhus, Denmark: Lægekredsforeningen for Aarhus Amt; 1979.
6. Morell DC, Gage HG, Robinson NA. Symptoms in general practice. J R Coll Gen Pract 1971;21:32–43.
7. Warndorff DK, Knottnerus JA, Huijnen LGJ, Starmans R. How well do general practitioners manage dyspepsia? J R Coll Gen Pract 1989;39:499–502.

8. Kristensen P, Sandbakken P, Johannessen T, Løge I, Hafstad PE. Fordøyelsessykdommer i almenpraksis. Tidsskr Nor Laegeforen 1985;105:728–31.
9. Hungin APS, Bramble MG, O'Callaghan H. Reasons for variations in the use of open access gastroscopy by general practitioners. Gut 1995;36:182–4.
10. Colin-Jones DG, Bloom B, Bodemar G. Management of dyspepsia: report of a working party. Lancet 1988;1:576–9.
11. Drossman DA, Grant Thompson W, Talley NJ, Funch-Jensen P, Janssens J, Whitehead WE. Identification of subgroups of functional gastrointestinal disorders. Gastroenterol Int 1990;3: 159–72.
12. Talley NJ, Zinsmeister AR, Schleck CD, Melton LJ. Dyspepsia and dyspepsia subgroups: a population-based study. Gastroenterology 1992;102:1259–68.
13. Agreus L, Svärdsudd K, Nyren O, Tibblin G. Irritable bowel syndrome and dyspepsia in the general population: overlap and lack of stability over time. Gastroenterology 1995;109:671–80.
14. Hansen JM Dyspepsia in primary care [thesis]. Odense, Denmark: University of Odense; 1996.
15. Kay L, Jørgensen T. Redefining abdominal syndromes. Results of a population-based study. Scand J Gastroenterol 1996;31: 469–75.
16. Kay L, Jørgensen T. Abdominal symptom associations in a longitudinal study. International J Epidemiol 1993;22:1093–100.
17. Meineche-Schmidt V, Krag E. Dyspepsia in general practice in Denmark. Scand J Prim Health Care. In press.
18. Armitage P, Berry G. Statistical methods in medical research. 3rd. ed. Oxford: Blackwell; 1994.
19. Meineche-Schmidt V, Krag E. Antisecretory therapy in 1017 patients with ulcerlike or refluxlike dyspepsia in general practice. Eur J Gen Pract 1997;3:125–30.
20. Bytzer P. Diagnosing dyspepsia: any controversies left? Gastroenterology 1996;110:302–6.
21. Johnsen R, Bernersen B, Straume B, Førde OH, Bostad L, Burhol PG. Prevalences of endoscopic and histological findings in subjects with and without dyspepsia. Br Med J 1991;302:749–52.

Received 19 February 1998
Accepted 7 October 1998

APPENDIX

The structure of the data was studied by using principal-components analysis (18). This method analyzes the pattern of correlation coefficients between the variables and combines groups of highly correlated variables into a smaller number of independent (uncorrelated) new variables explaining a large part of the variation (18).

Primarily, we extracted principal components with eigenvalues higher than 1 (Kaiser's criterion). In this manner we obtained six principal components. However, the last two components had eigenvalues only slightly higher than 1 (1.10 and 1.06, respectively) and did not provide additional meaningful information. Hence we only extracted the first four principal components, accounting for 36.1% of the variance. Normalized varimax rotation was performed to yield components easier to interpret. In the analysis the symptom variables were scored 1 for present or 0 for absent.

Using the coefficients of the obtained principal components we calculated the score of each patient for each of the components, using standard methods (18). The distribution of the patient scores for each principal component was studied to ascertain whether the patients showed signs of clustering. One-, two-, and three-dimensional plots were studied, each dimension corresponding to a principal component.

Calculation of principal components factor scores for individual patients.

Corresponding to the 4 identified principal components, 4 new independent (uncorrelated) variables can be made from the original 22 variables included in the analysis. These new variables were calculated by using the factor score coefficients for the principal components. These are obtained as a part of the principal-components analysis and are presented in Appendix Table I together with the mean and standard deviation of the original variables. The mean and standard deviation are necessary because in principal-components analysis the variables are being analyzed in their standardized form—that is, after subtraction of the mean and subsequent division by the standard deviation; thus standardized variables have a mean of zero and a standard deviation of one.

Each new factor score variable Y_j is calculated as follows:

$$Y_j = b_{j1}(x_1 - M_1)/SD_1 + ... + b_{ji}(x_i - M_i)/SD_i + ... + b_{jp}(x_p - M_p)/SD_p, \quad \text{[Equation 1]}$$

where b_{ji} is the factor score coefficient for the factor score Y_j to be used for the original variable x_i with mean value M_i and standard deviation SD_i, j being 1, 2, 3, or 4 (corresponding to the four new factor score variables) and i being 1, 2, 3 ... p (corresponding to the p (= 22) original variables).

The distribution of the four factor score variables in the total sample of 7270 patients is shown in Fig. 1. The distributions in Fig. 1 all have a mean value of zero and a standard deviation of one. For a new patient the four factor scores can be calculated from the patient variables and the information in Appendix Table I, using equation 1 shown above. Such calculation is rather tedious. To facilitate clinical use in new patients, a chart for easy approximate calculation is presented in Table IV. That chart is based on the information in Appendix Table I, with each of the terms in equation 1 being precalculated for each symptom (being present minus being absent) multiplied by 2 and rounded to integers. The correction terms in the chart are adjustments to keep the mean of the factor score distributions at zero, just as in Fig. 1, but with a double standard deviation. The reasons for choosing the chart to present only approximate figures in the form of small integers are to facilitate clinical usage and to emphasize that results obtained by the principal-components analysis are not precise but statistical approximations that fit the present data but may not necessarily be fully applicable to other patient populations.

Appendix Table I. Principal-components factor score coefficients, means and standard deviations for the component variables

Variable	Component				Mean	s*
	A	B	C	D		
Epigastric pain	0.2511	0.0320	0.0728	−0.0056	0.6310	0.4826
Pain relieved by antacids	0.2430	0.0729	−0.0772	−0.0771	0.3724	0.4835
Acid regurgitation	0.2382	0.0663	−0.0064	0.0817	0.4171	0.4931
Other abdominal pain	−0.2246	−0.0356	−0.0145	−0.0341	0.2535	0.4351
Heartburn	0.1842	0.0013	−0.1029	0.2124	0.2622	0.4399
Pain in the night	0.2011	0.0746	0.0921	−0.0602	0.4135	0.4925
Pain relieved by food	0.1780	0.0884	0.0709	−0.3790	0.2912	0.4544
Bloating	0.0754	0.3609	−0.0367	−0.0200	0.4198	0.4936
Pain relieved by stools or flatus	0.0136	0.3139	−0.0600	−0.0468	0.2519	0.4341
Constipation	0.1048	0.3165	−0.0780	0.0425	0.1656	0.3718
Incomplete evacuation	0.1162	0.3068	0.0440	−0.0187	0.0827	0.2754
Upper abdominal pain	−0.0592	0.1892	0.0085	0.0104	0.3585	0.4796
Loose stools	−0.0121	0.1691	0.1193	−0.0911	0.2344	0.4236
Female sex	0.0139	0.0972	0.0359	0.0967	0.5638	0.4959
Nausea	0.0390	0.0779	0.3949	0.0928	0.3165	0.4651
Morning vomiting	−0.0060	−0.0610	0.4096	0.0165	0.0468	0.2112
Pain relieved by vomiting	0.0964	−0.0042	0.3649	0.0440	0.1315	0.3380
Pain in the morning	0.0112	0.0571	0.2500	−0.0702	0.1664	0.3725
Age	0.1453	0.1548	−0.2395	0.1149	47.115	18.691
Pain after meals	0.0532	0.1079	−0.0457	0.4884	0.3625	0.4807
Dysphagia	0.0333	−0.0461	−0.0352	0.4268	0.0337	0.1805
Weight loss	0.0073	−0.0320	0.1402	0.2892	0.0388	0.1931

* s = standard deviation.