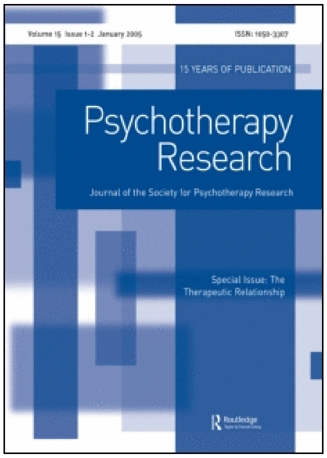

Psychotherapy Research
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713663589

An individual differences perspective on change in psychotherapy: The case of health care utilization

A. Lazar [a]; R. Sandell [b]; J. Grant [b]
[a] Stockholm County Council Institute of Psychotherapy, Stockholm, Sweden
[b] Department of Behavioural Sciences, Linköping University, Stockholm, Sweden

Online Publication Date: 01 November 2007
To cite this Article: Lazar, A., Sandell, R. and Grant, J. (2007) 'An individual differences perspective on change in psychotherapy: The case of health care utilization', Psychotherapy Research, 17:6, 690 - 705
To link to this article: DOI: 10.1080/10503300701275310
URL: http://dx.doi.org/10.1080/10503300701275310

Psychotherapy Research, November 2007; 17(6): 690–705

An individual differences perspective on change in psychotherapy: The case of health care utilization

A. LAZAR[1], R. SANDELL[2], & J. GRANT[2]

[1]*Stockholm County Council Institute of Psychotherapy, Stockholm, Sweden,* [2]*Department of Behavioural Sciences, Linköping University, Stockholm, Sweden*

(Received 7 June 2006; revised 7 February 2007; accepted 12 February 2007)

Abstract
The purpose of this study was to explore systematic interindividual variation in change of a number of health care utilization variables (HCUVs) during psychotherapy and identify patient characteristics associated with this variation. Three-wave panel data from 420 patients were analyzed with nonparametric latent class regression followed by chi-square interaction analyses among patient variables. For the various HCUVs, three to six classes were identified, with widely different patterns of change during treatment. Axis I diagnosis, chronicity, functional impairment, gender, and level of education were among the patient characteristics that differentiated the classes. It was concluded that main effects analyses seriously distort heterogeneity of change and that health care utilization, unless it is a specific therapeutic aim, may be irrelevant as an indication of outcome of psychotherapy.

There is evidence that psychological interventions—psychotherapy proper or mental health treatment (Mumford, Schlesinger, Glass, Patrick, & Cuerdon, 1984)—may contribute to substantial medical offset (Beutel, Rasting, Stuhr, Rüger, & Leuzinger-Bohleber, 2004; Breyer, Heinzel, & Klein, 1997; Chiles, Lambert, & Hatch, 1999; Dührssen, 1962; Dührssen & Jorswieck, 1965; Follette & Cummings, 1967; Gabbard, Lazar, Hornberger, & Spiegel, 1997; Keller, Westhoff, Dilg, Rohner, & Studt, 2001; Mumford et al., 1984). According to Chiles et al. (1999), however, this effect is contingent on whether the treatment has this specific aim or not. The evidence of side effects of "ordinary" psychotherapy on health care appears to be weaker.

Accordingly, Lazar, Sandell, and Grant (2006) found no significant change on a number of health care utilization variables (HCUVs) among patients in long-term psychodynamic therapy despite robust positive change on subjective health measures. As the authors noted, a number of economic and social–psychological complications might explain the negative findings. The major share of health care costs in Sweden is covered by taxes and national insurance, and people's criteria for using health care and insurance seem to have developed in a more liberal direction, both from the point of view of the medical doctors and of the patients in general.

Consequently, the reasonable relation between subjective health and health care utilization (HCU) has been offset by social–psychological factors, to the great concern of the national authorities. Thus, in a further study, Lazar, Sandell, and Grant (in press) found only weak relations between the two. Autoregressive structural models found great stability in subjective health across 3 years and moderate stability in HCU, varying widely among different variables, but only a few, weak, inconsistent, and uninterpretable links between the two.

However, common sense and experience suggest that people's criteria differ for declaring themselves ill or in need of treatment. Some people may be objectively ill but continue working nevertheless, whereas others may report sick with a slight headache. Some will feel ill without any corroboration in medical tests, whereas others will not do so despite serious diagnoses. Therefore, the relations between being ill, feeling ill, and acting on any of the two will vary depending on personal factors. In addition, the effect of psychotherapy on patients' HCU may likewise vary depending on individual differences in values, attitudes, and, certainly, objective bodily processes and external circumstances. As a consequence, looking for general effects of psychotherapy on HCU may be a mistaken and futile approach, particularly in view of the suggestion that the patient

Correspondence: Rolf Sandell, Department of Behavioural Sciences, Linköping University, Fredrikshovsgatan 3A, Stockholm SE-115 23, Sweden. E-mail: rolsa@ibv.liu.se

ISSN 1050-3307 print/ISSN 1468-4381 online
DOI: 10.1080/10503300701275310

factor accounts for the dominant share of treatment outcome in general (Clarkin & Levy, 2004; Lambert, 1992; Norcross & Lambert, 2006; Wampold, 2001).

In this study, we explored the development or change in a number of HCUVs across stages of psychotherapy through the lens of latent class (LC) regression analysis (Bouwmeester, Sijtsma, & Vermunt, 2004; Vermunt & van Dijk, 2001) with the aim of identifying subgroups of patients differentially affected by long-term psychodynamic therapy. The data derive from the Stockholm Outcome of Psychotherapy and Psychoanalysis Project (STOPPP).

Based on our previous observations of weak relations between HCU and subjective health and the substantial variation in both kinds of variables, the general hypothesis tested was that there was systematic heterogeneity among patients in terms of change in HCU during treatment and follow-up and that this heterogeneity is partly accounted for by demographic and psychiatric pretreatment variables.

Method

Design

The design has been extensively described by Lazar et al. (2006), Blomberg, Lazar, and Sandell (2001), and Sandell et al. (2000). It was a quasi-experimental, partly cross-sectional, partly longitudinal design based on a postal three-wave panel survey among patients in psychotherapy or psychoanalysis and a survey among their therapists.

Procedure

A sample of 756 persons in subsidized treatment or on the waiting list for such subsidization was selected so as to ensure that it consisted of people who had terminated their treatments as well as people who were in ongoing treatment or who had not yet started treatment.

The Well-Being Questionnaire (WBQ), which included a number of self-rating scales, was distributed to all 756 persons in 1994 and again in 1995 and 1996 to all who had responded the first year. Return rates of 78%, 86%, and 88%, respectively, produced a panel of 445 persons (59%), each with three observations.

With three possible treatment states (pretreatment, in treatment, posttreatment) and three panel waves, it was possible to establish an ordinal time scale with nine successive steps corresponding to stages in treatment: three before treatment, three during treatment, and three after treatment. We located each patient in the panel each year on this treatment stage scale. Of the 445 respondents, 344 were in psychotherapy, 76 were in psychoanalysis, and 13 were in various low-dose therapies; 12 never commenced treatment, which reduced the stage scale to eight steps. We excluded the last two groups,[1] leaving a total sample of 420 patients distributed across the stage scale.

The basic model for our analyses was nonparametric LC regression analysis for clustering the patients on the basis of their outcome trajectories on various measures of HCU.

Patients

Patient data are presented in Table I. The typical patient was female (77%), unmarried (58%) or divorced (17%), with children (52%). The majority (79%) had some college or university education and typically worked in the health care, education, or social sectors. The mean age was 38.7 years ($SD = 8.3$). More than half of the patients (58%) had at least one *Diagnostic and Statistical Manual of Mental Disorders* (fourth edition; *DSM–IV*) Axis I diagnosis (typically mood or anxiety syndromes), 12% had an Axis II personality disorder (5% of the patients had both a personality disorder and at least one Axis I diagnosis), and V codes were assigned to 33% of the patients. The average Global Assessment of Functioning (GAF; *DSM–IV*, Axis V; American Psychiatric Association, 1994) score ($M = 59.7$, $SD = 5.5$) indicates a moderate dysfunction level. More than half of the patients (66%) had been in psychotherapy previously. Further details on the patient sample are given in Blomberg et al. (2001).

Treatments and Therapists

The treatments were formally defined as psychotherapy once- or twice-a-week treatment ($n = 344$; 82%) or psychoanalysis ($n = 76$; 18%) three to five times a week, in accordance with the specification on the referrals. Both kinds were planned to be long term, according to the referrals, and all were individual treatments. The treatments were not manualized or standardized with respect to duration, session frequency, technique, and so on. The two treatment forms were not separated in this study.

The 294 therapists with patients in the sample were licensed by the National Board of Health and Social Welfare, and some were also trained and members of one of the two psychoanalytic societies in Sweden. Of the therapists, 95% claimed to be "rather strongly" or "strongly" oriented toward a psychoanalytic or psychodynamic theoretical position, and 11% claimed also to share "strongly" or "rather strongly" an eclectic position. Further details

Table I. Patient Characteristics

Variable	%	*M*	*SD*
Sociodemographic characteristics			
Male	23		
Age (years)		38.7	8.3
Married or divorced	42		
Has children	52		
College education	79		
Swedish origin	95		
Psychiatric characteristics			
DSM–IV Axis I diagnosis	58		
DSM–IV Axis II diagnosis	12		
GAF-C		59.7	5.5
GAF-L		52.2	10.8
VIS		1.3	1.5
No. problem years		2.7	1.2
Health care utilization			
No. of somatic consultations		2.3	3.1
No. of psychiatric consultations		0.7	3.9
No. weeks somatic inpatient		0.2	1.0
No. weeks psychiatric inpatient		0.3	2.1
Level of medicine consumption[a]		2.3	1.2
No. days absent from work		30.5	68.2

Note. *DSM–IV* = *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.); GAF-C = Global Assessment of Functioning Scale, current; GAF-L = Global Assessment of Functioning Scale, lowest level after age 18; VIS = Vocational Impairment Scale.
[a]Rated on a 0 to 4 scale.

on the treatment providers are given in Sandell et al. (2006a, 2006b).

Assessment Procedures

The WBQ contained the following sections, with standard items or questions on (a) demography and socioeconomy and familial, vocational, and financial situation; (b) ongoing psychotherapy; (c) previous treatments, including psychotherapy, for psychological distress; (d) current health status and health care utilization in the past 12 months; (e) severity of current and prior psychological problems; and (f) occupational activities (including studies) during the past 12 months. In addition, the following three self-rating instruments were included: Symptom Checklist-90 (SCL-90; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974); Social Adjustment Scale (SAS; Weissman & Bothwell, 1976; Weissman, Prusoff, Thompson, Harding & Myers, 1978), in a revised version to suit Swedish users in the 1990s; and Sense of Coherence Scale (SOCS; Antonovsky, 1987). These scales were not used in this study.

Patients' pretreatment status. Demographic information (sex, age, marital status, number of children, and level of education) was collected principally from the WBQ. Another item in the WBQ gave the number of years with the current psychiatric problems according to the patient.

Various diagnostic indicators were assessed on the basis of the referrals. Thus, each patient was grossly diagnosed for the presence of *DSM–IV* Axis I or II disorders (American Psychiatric Association, 1987) before treatment. Patients were also assessed on the GAF scale. In addition to the current GAF score (GAF-C), a rating was made of the lowest level of functioning after age 18 years (GAF-L). Also, a specially designed Vocational Impairment Scale (VIS) rated vocational impairment on a 5-point scale ranging from 0 (*no impairment*) to 5 (*totally unable to work for more than year*). Interrater agreement (intraclass correlations), tested with three raters (all licensed psychologists) on 20 referrals, were, .69 for Axis I diagnoses, .51 for Axis II diagnoses, .69 and .88 for the GAF-C and the GAF-L, respectively, and .80 for the VIS.

Patient outcome measures. The following HCUVs were assessed by patients' self-reports in the WBQ:

1. Number of consultations with medical doctors: "During the past 12 months, have you visited with a medical doctor (MD) for somatic (bodily) diseases or troubles of your own? If so, approximately how many times totally during the past 12 months?"
2. Number of consultations with psychiatrists: "During the past 12 months, have you visited an MD for psychological troubles or problems of your own? If so, approximately how many times totally during the past 12 months?"
3. Number of weeks of inpatient treatment in hospital, nursing home, and so on for somatic ailments: "During the past 12 months, have you been admitted as an inpatient to a hospital or any kind of treatment institution for somatic (bodily) diseases or troubles? If so, approximately how many weeks total during the past 12 months?"
4. Number of weeks of inpatient treatment in hospital, nursing home, or other for psychiatric problems: "During the past 12 months, have you been admitted as an inpatient to a hospital or any kind of treatment institution for psychological troubles or problems? If so, approximately how many weeks total during the past 12 months?"
5. Level of medicine consumption: "During the past 12 months, have you been using any kind of medicine for somatic (bodily) or psychological troubles?" Responses were based on a 5-point scale (0 = *not at all,* 4 = *regular consumption of*

several drugs over the whole year). Types of medical drugs were not differentiated.

6. Number of days of absence from work or other regular occupation as a result of ill health during the past 12 months, including disability pension: "During the past 12 months, how long have you been sick-listed or absent from work or your regular occupation (e.g., studies) on account of sickness of your own, in total?" Responses were based on a 4-point scale (1 = *not at all,* 4 = *more than 1 month*). Respondents were asked, on each step if applicable, to give the number of days, weeks, and months absent. Psychiatric and somatic causes were not differentiated. The information given was converted into number of days: 1 week was counted as 5 days and 1 month as 21.66 days.

Norm Group

To establish a standard for evaluating patient outcome in relation to normality, the WBQ was also distributed in two nonclinical groups: (a) a random community sample of 400 persons between 20 and 69 years of age in Stockholm County (obtained through the services of the National Post) and (b) a sample of 90 psychology students. The psychology students were an introductory class outside the national professional psychologists training program. The norm groups responded to the questionnaire only once, in May 1994. Without any reminders, the response rates in the two groups were 37% and 76%, respectively. The responders in the two groups had almost identical mean values on the self-rating scales, and they were, therefore, collapsed into one group. The pooled norm groups of respondents ($N = 214$) were demographically rather similar to the patient sample, with a majority of women (63 vs. 77%), single (42 vs. 58%) or divorced (15 vs. 17%), of Swedish origin (93 vs. 95%), and with at least some college or university training (55 vs. 79%). Their mean age was 33.9 years ($SD = 10.1$) versus 38.7 years ($SD = 8.3$). The group was used merely for descriptive comparison purposes.

Statistical Analyses

Nonparametric LC regression modeling with repeated measures (Vermunt & van Dijk, 2001) was used to analyze the HCUVs. LC analysis (LCA) is a statistical method for finding subtypes of cases, so-called latent classes, in a multivariate data set. LCA is used analogously to cluster analysis. That is, given a sample of cases measured on several variables, one wishes to know whether there are a limited number of subgroups (types, clusters, categories) into which the cases fall. In this study, the cases are patients and the variables are repeated measurements of HCUVs. These repeated measurements are regressed, patient by patient, on a temporal variable, and the patients are grouped on the basis of similarity of the regression parameters, which function as treatment outcome parameters. The subgroups found are called *latent classes*, "latent" because class membership is not an observable variable. When the corresponding empirical observations are analyzed, the classes are referred to as *observed groups*.

Nonparametric LC models are less subject to biases as a result of violations of conventional assumptions about linearity, normality, homoscedasticity, independence, and homogeneity. An LC model introduces a latent nominal variable for classes or clusters. This class variable serves as a moderator in interaction with one or several observed predictors, treatment stage in this case. Typically, LC regression analysis does four things simultaneously: (a) identifies latent classes, (b) estimates regression models for each class, (c) tests covariates to predict class membership, and (d) assigns cases to classes. When the dependent variable is a repeated measure, LC regression may be seen as a case of multilevel modeling.

We used the Latent GOLD© 4.0 (LG) software (Vermunt & Magidson, 2005) and analyzed the HCUVs with patients as the units of analysis, each patient with three observations 1 year apart along the treatment stage scale. Complete data were available for 420 patients with 1,260 observations spread across the stage scale. All HCUVs except medicine consumption were specified as Poisson count variables in estimating the models; medicine consumption was specified as an ordinal categorical variable.

A regression model for each class describes the development of the assigned patients in terms of a regression coefficient, representing these patients' average rate of change across stages, and an intercept, representing their mean pretreatment state as measured by the HCUV. The number of classes was determined by the minimum Bayesian information criterion (BIC), based on the log-likelihood and considering the degrees of freedom. (In one case in which the BIC suggested an inordinate number of classes, we let a scree-type criterion applied to the squared multiple correlation values determine the solution). Following this, a number of covariates were introduced: sex, age (in one of five categories), marital status, number of children, education, number of years with psychiatric problems, Axis I diagnosis, Axis II diagnosis, GAF-C, GAF-L, and VIS. The two GAF variables were transformed into seven categories in as rectangular a distribution as possible, using a LG subroutine. Finally, we introduced patient's treatment form as another covariate. The covariates were specified as inactive, meaning that they were not allowed to affect the classifications. After specifying the LC model, the associations between the LC variable and the covariates were fed into a chi-square automatic interaction detector (CHAID) algorithm (SI-CHAID© 4.0; Magidson, 2005). Based on the chi-square statistic, controlling for the associations among the covariates, SI-CHAID explores, in a stepwise procedure, the extent to which the covariates, in combination, may predict membership in the latent classes. The result is visualized in a tree structure accounting for as much of the between-classes differences as possible, given the associations between the covariates and the chosen significance level, $p < .05$.

Results

Preliminary Tests of the Design

To be able to interpret the regression coefficients in terms of patients' change as a function of treatment stages, it is of vital importance that the stage scale was not confounded with other variables. We, therefore, explored the associations between the stage scale and a number of variables pertaining to the therapists, the patients, and the treatments. Testing more than 30 variables for their correlations with time, we found only one with a near-significant correlation: patients' number of previous treatments in psychiatric open care ($-.10$, $p = .055$). We concluded that our stage scale was free of obvious strong confounds.

The scale properties of the HCUVs were also explored. Across all 1,260 observations, all HCUVs had strong positive skewness; the number of observations at the minimum level varied from 27% (absence days) to 97% (psychiatric inpatient weeks).

The output of LG is copious, with multitudinous parameters and tests of significance for each class. In order not to unduly burden the following account, significance tests are reported selectively, and the results are principally given by graphs. For the interested reader, technical output may be obtained from Rolf Sandell.

Number of Consultations

Somatic consultations. To introduce the newcomer to LC regression analysis, we provide a detailed account of these first results. The undivided group had a significant decreasing trend ($\beta_0 = 0.873$, $\beta_1 = -0.038$, $z_{\beta 1} = 3.36$, $p = .001$).[2] This corresponds to a decrease in the mean number of consultations from 2.39 the year before treatment to 1.91 the third year after termination, an average stagewise reduction by 0.08 visits with an MD for somatic ailments. The change across treatment stages is shown in Figure 1 (top panel). A selection of the LG output is displayed in Tables II and III. The LC regression analysis suggested a six-class solution according to the minimum BIC. This solution accounted for 68% of the total variance ($R^2 = .677$), whereas the one-class solution accounted for less than 1%. A graph of the modeled changes of the six latent classes across the stage scale is given in Figure 1 (middle panel). For demonstrative purposes, the corresponding observed changes are graphed in the bottom panel of Figure 1. The classes are ordered according to size and are named with acronyms in an attempt to describe the change pattern (initial and final levels) according to the model.[3] The mean in the norm group (1.24 visits) is indicated by a horizontal line, and another line, 1.28 *SD* above the norm group mean, indicates the division between the 90% of the norm group with the lowest number of visits and the 10% with the highest (4.25 visits). This was taken, arbitrarily, as the division between a subclinical and a clinical subgroup in an unselected sample and is called the caseness criterion, following Derogatis and Lazarus (1994).

The largest class, LVL (low to very low), comprising 45% of the patients, began their treatment with a mean of nearly three visits. The norm group had between 1 and 1.5 visits. During the treatment, the patients lowered their number of visits significantly ($\beta_0 = 1.048$, $\beta_1 = -0.137$, $z_{\beta 1} = -4.86$, $p < .001$) to the mean level of the norm group. This corresponds to a reduction from 2.85 to 1.25 visits, that is, an average reduction of 0.27 per stage, more than three times higher than in the undivided sample (which, of course, included class LVL). The next largest class, VLZ (very low to zero), with 25% of the patients, also decreased their visits significantly, to a level close to 0, and class MM (medium to medium; 14%) also had a significant downward trend, from a higher initial level to the clinical/nonclinical division line. In contrast to these groups, class ZM (zero to medium; 11%) had a rather sharp increase in visits starting late in treatment and accelerating after its termination, from close to 0 to slightly more than 3. Even more dramatic a development was shown by class VLVH (very low to very high; 3%), with an apparently ever-rising number of visits, from less than one to more than 18 visits the third year after termination. Finally, class VHH (very high to high; 2%; eight patients) showed a clearly declining trend, albeit not significant, on account of the small class size. The CHAID analysis revealed no covariate with a significant association with the classification.

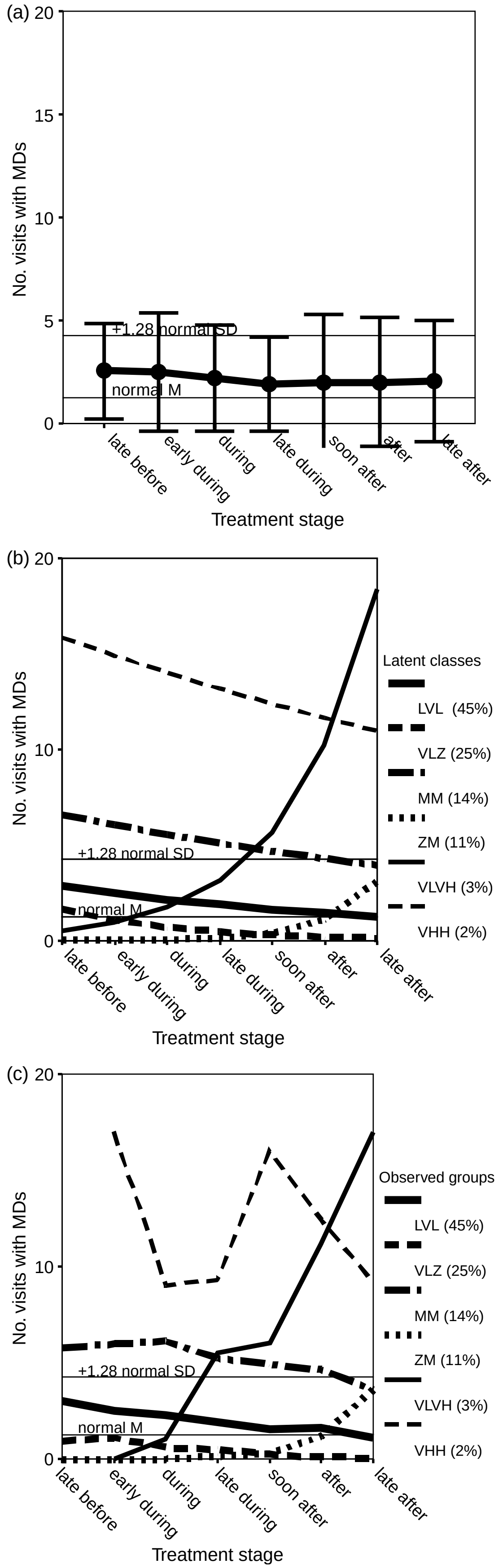

Figure 1 (Continued)

Table II. Selected LC Regression Analysis Output to Determine Number of Latent Classes Relating Number of Consultations With Medical Doctors to Treatment Stages

Model	No. classes	BIC	No. parameters	R^2
1	1	6001.94	2	.00
2	2	4922.14	5	.44
3	3	4735.86	8	.56
4	4	4671.89	11	.62
5	5	4662.42	14	.65
6	6	4661.13	17	.68
7	7	4667.11	20	.69

Note. The six-class solution was selected on the basis of minimum BIC. LC = latent class; BIC = Bayesian information criterion.

Psychiatric consultations. After excluding one outlier patient, the nondivided sample had a nonsignificant decrease in number of consultations ($\beta_0 = 0.392$, $\beta_1 = -0.030$, $z_{\beta 1} = -1.38$, $p = .16$). The optimal number of latent classes, according to the BIC, was six. Their modeled profiles are displayed in Figure 2 ($R^2 = .68$). (The norm group mean was 0.08 visits and the caseness criterion 0.85; neither is shown in the figure because both are too close to zero to be clearly visible.) Obviously, the largest class, ZZ (zero to zero), including 70% of the sample, had virtually zero visits initially and remained at a low level, although it had a small, yet significant, increase to 0.10. A sharp and significant decline toward zero level as treatment began was evident in class MZ (medium to zero; 10%). Together with class MVL (medium to very low), also with a significant declining trend, and class MM, with a nonsignificant increase, these three classes with negative or zero trends had more than 85% of the sample. In contrast, there were two classes (ZM and VLVH) with significant increasing trends.

The CHAID analysis revealed that four of the covariates significantly (ps $< .029$) predicted class membership: VIS, GAF-L, GAF-C, and number of problem years.[4] Because of their interrelations, only three made unique significant contributions to class prediction, and the resulting tree diagram is displayed in Figure 3. The first division of the entire sample was between patients with high ($n = 59$) and low ($n = 360$) VIS scores. Those with

Figure 1. Observed trajectory for number of medical consultations for the entire sample across treatment stages (upper panel; ± 1 *SD* indicated by wings); model trajectories for the six latent classes (middle panel); and observed trajectories for the corresponding six groups of patients (lower panel). The classes and groups are placed in order of size and are named with acronyms to describe the approximate change pattern (initial level and final level). (H = high; L = low; M = medium; VH = very high; VL = very low; Z = [close to] zero).

Table III. Characteristics of the LC Regression Analysis-Derived Classes Relating Number of Consultations With Medical Doctors to Treatment Stages

	Class						
Variable	1	2	3	4	5	6	Overall
Class size (%)	45	25	14	11	3	2	
R^2	.09	.19	.05	.59	.74	.05	.65
β parameters							Wald
Intercept (β_0)	$1.05^{\dagger\dagger\dagger}$	$0.49^{\dagger}$	$1.89^{\dagger\dagger\dagger}$	$-5.01^{\dagger\dagger\dagger}$	−0.62	$2.76^{\dagger\dagger\dagger}$	528.68***
Slope (β_1)	−0.137***	−0.414***	−0.086***	1.024***	0.589***	−0.062	82.12***

$^{\dagger}p<.05$; $^{\dagger\dagger\dagger}p>.001$ (one-tailed tests).
***$p>.001$; **$p<.01$; *$p<.05$; (*) $p<.10$ (two-tailed tests).

low VIS scores were further divided according to the number of problem years and those with high VIS according to their GAF-L scores. Thus, four covariate groups (*segments*) were obtained. The largest (segment A), with 252 of the 419 patients (60%), had a slight overrepresentation of ZZ patients, with no or almost no psychiatric visits, and slight underrepresentation of patients from classes ZM, MVL, VLVH, and MM. These were the persons with no to moderate vocational impairment and 3 or fewer years of problems. The next segment (B) had 108 patients (26%), among whom ZM patients were more frequent than expected and class ZZ patients less so. Patients in this segment had no to moderate vocational impairment but a history of 4 or more years with psychological problems as they began their treatment. Another segment, with 40 patients (C; 10%), with severe to complete vocational impairment and low GAF-L (<53.5), had a striking underrepresentation of ZZ patients and an overrepresentation of classes MVL, VLVH, and MM and, not as dramatically, of ZM patients. Finally, a small segment of 19 patients (D; 5%) had severe to complete vocational impairment but relatively high GAF-L (>53.5). This segment was heavily dominated by ZZ patients.

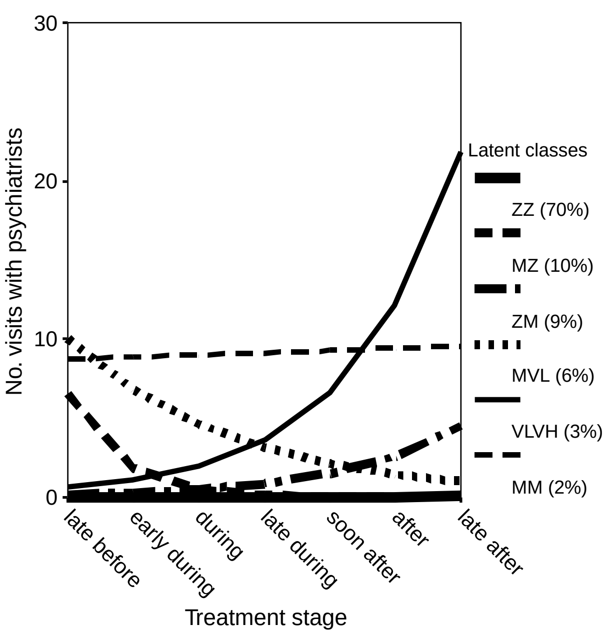

Figure 2. Outcome trajectories for number of psychiatric consultations for the six latent classes model. (H = high; L = low; M = medium; VH = very high; VL = very low; Z = [close to] zero).

Hospital Treatment

Somatic hospitalization. In the entire group, there was a small but significant increase ($\beta_0 = -2.118$, $\beta_1 = 0.093$, $z_{\beta 1} = 2.23$, $p = .026$) at a very low level, from 0.12 the year before treatment to 0.21 weeks the third year after termination, an average increase of 0.015. Three latent classes ($R^2 = .28$) were estimated, and the model profiles are shown in Figure 4 (upper panel). The largest class, ZZ, had 69% of the patients and a nonsignificant decrease in hospitalization, from 0.03 to 0.01 weeks. Class VLM (very low to medium) increased from 0.17 to 0.39, which was also nonsignificant, and so was the increase in class VHVH (very high to very high), from 2.72 to 2.90 weeks. None of the covariates predicted class membership.

Psychiatric hospitalization. One patient was excluded from this analysis as an outlier. The entire sample increased their number of inpatient weeks significantly, from 0.14 the year before to 0.35 the third year after treatment ($\beta_0 = -1.956$, $\beta_1 = 0.151$, $z_{\beta 1} = 4.31$, $p < .0001$). Four latent classes were extracted on the basis of the BIC, and, paradoxically, all had nonsignificant reductions in hospital weeks ($R^2 = .56$).[5] The profiles of the classes are shown in Figure 4 (lower panel; norm group mean is not shown because it was too close to 0). The largest class, ZZ, with 94% of patients, reduced their number of weeks from 0.01 to 0.00, whereas the other three classes had progressively greater reductions at progressively higher levels. Three covariates had significant (ps < .026) associations with the classification, GAF-L, GAF-C, and treatment type.

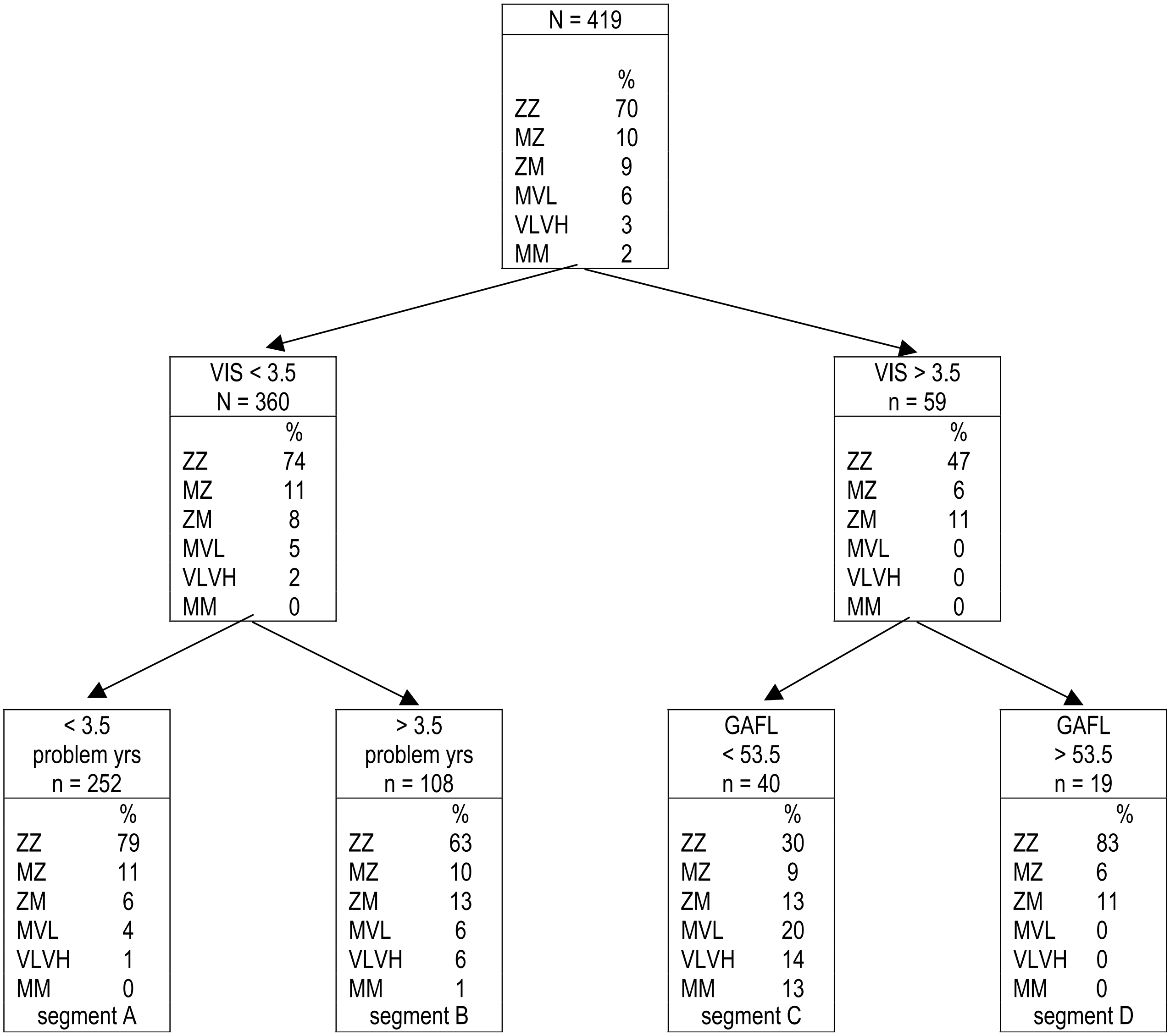

Figure 3. SI-CHAID-generated tree diagram relating the six latent psychiatric consultations classes to the nonredundant covariates. (VIS = Vocational Impairment Scale; GAF-L = Global Assessment of Functioning Scale, lowest level after age 18. Specific segments refer to groups of combinations of covariates; H = high; L = low; M = medium; VH = very high; VL = very low; Z = [close to] zero).

Only GAF-L made an independent contribution to class prediction. The division was at a GAF-L score of 53.5; those with higher ratings were dominated by ZZ patients, and there was an overrepresentation of lower GAF-L levels in the other classes.

Medicine Consumption

In the undivided sample, there was a low and nonsignificant decrease ($\beta_1 = -0.0139$).[6] Translated into raw scores, this is a reduction from 2.36 to 2.22 on the 5-point scale, a 0.02 change per stage. Again, the four-class solution yielded the lowest BIC value ($R^2 = .72$). The changes were nonsignificant in all classes. HH (high to high) patients increased their consumption the most, from 2.67 to 3.24 on the 5-point scale. Although this might have been a significant change in a larger sample, the four classes represented essentially stable, or treatment-independent, levels of medicine consumption. The model trajectories are shown in Figure 5.

Five covariates were significantly (ps $< .04$) associated with the classification: GAF-C, GAF-L, sex, Axis I diagnosis, and VIS. Only the GAF-C was uniquely useful in the prediction. Persons with GAF-C less than 57.5 were overrepresented in classes VHVH and HH, and persons with higher GAF-C were overrepresented in classes LL (low to low) and MM.

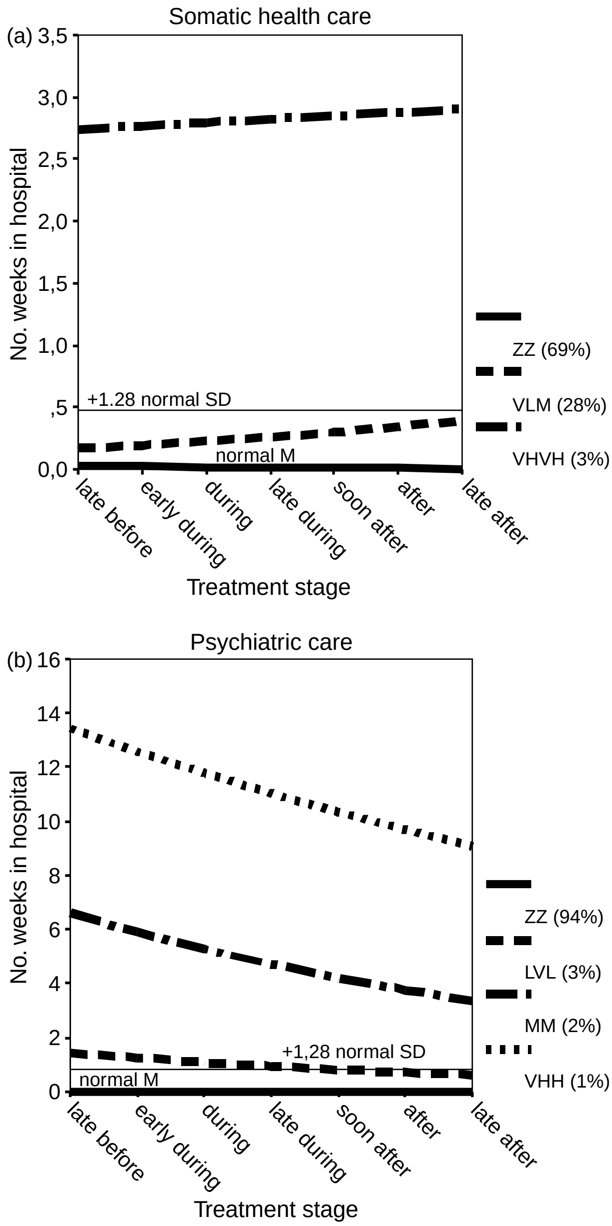

Figure 4. Outcome trajectories for number of hospital weeks in somatic care for the three latent classes model (upper panel) and in psychiatric care for the four latent classes (lower panel). (H = high; L = low; M = medium; VH = very high; VL = very low; Z = zero [or close to zero]).

Inability to Work

For this particular analysis, we used a subgroup of patients whom we defined as belonging to the labor force (Lazar et al., 2006). Based on all available information about each patient's occupational status (e.g., type of occupation, percentage of full-time job, possible reasons for not being in full-time employment) at each panel wave, we included persons if they had had at least 75% of full-time employment all 3 panel years, whether or not they had been on disability pension or temporary disability pension or had been sick-listed any time during the last 12 months. This group included not only employees but also self-employed persons and freelancers (e.g., in artistic professions). The group consisted of 216 persons. On the basis of the WBQ, we calculated the number of days of absence from work or other occupation as a result of ill health during the past 12 months (psychiatric and somatic causes were not differentiated).

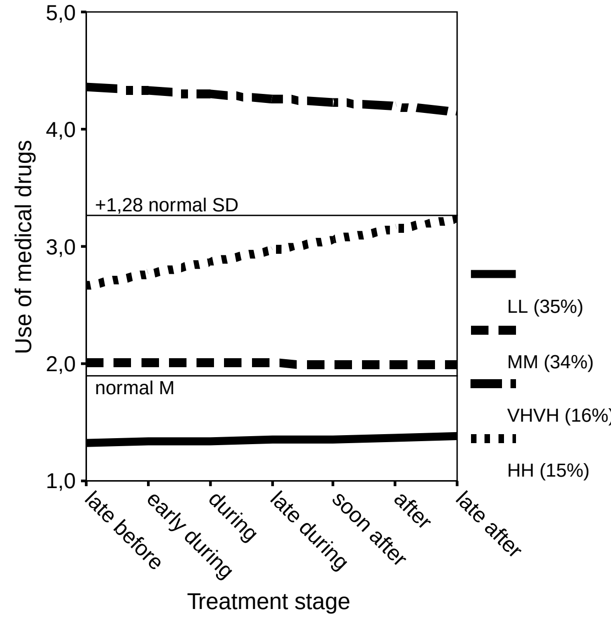

Figure 5. Outcome trajectories for medicine consumption (on the 5-point rating scale) for the four latent classes model. (H = high; L = low; M = medium; VH = very high; VL = very low; Z = [close to] zero).

When we analyzed this variable, the whole group changed from 45.45 to 39.78 days, a significant reduction ($\beta_0 = 3.817$, $\beta_1 = -0.022$, $z_{\beta 1} = -6.19$, $p < .001$). The minimum BIC was reached with 13 classes. This indicates very high heterogeneity in the sample. Because the squared multiple correlation value increased by less than .01 after the fifth latent class, we chose a solution with six classes ($R^2 = .95$). The profiles of the classes are shown in Figure 6.

Class VLZ had almost 33% of the sample and showed a significant downward trend in days ($\beta_0 = 1.622$, $\beta_1 = -0.263$, $z_{\beta 1} = -4.35$, $p < .001$), from slightly over 5 days to slightly over 1. In comparison, the norm group had a mean of 10 days ($SD = 32$). The reductions in classes VLVL (very low to very low) and VHVH were not significant. The upward trends in classes ZM, LM (low to medium), and MVH (medium to very high) were all significant, and these classes comprised 25% of the sample.

The following five covariates had significant chi-square associations with the classification, again

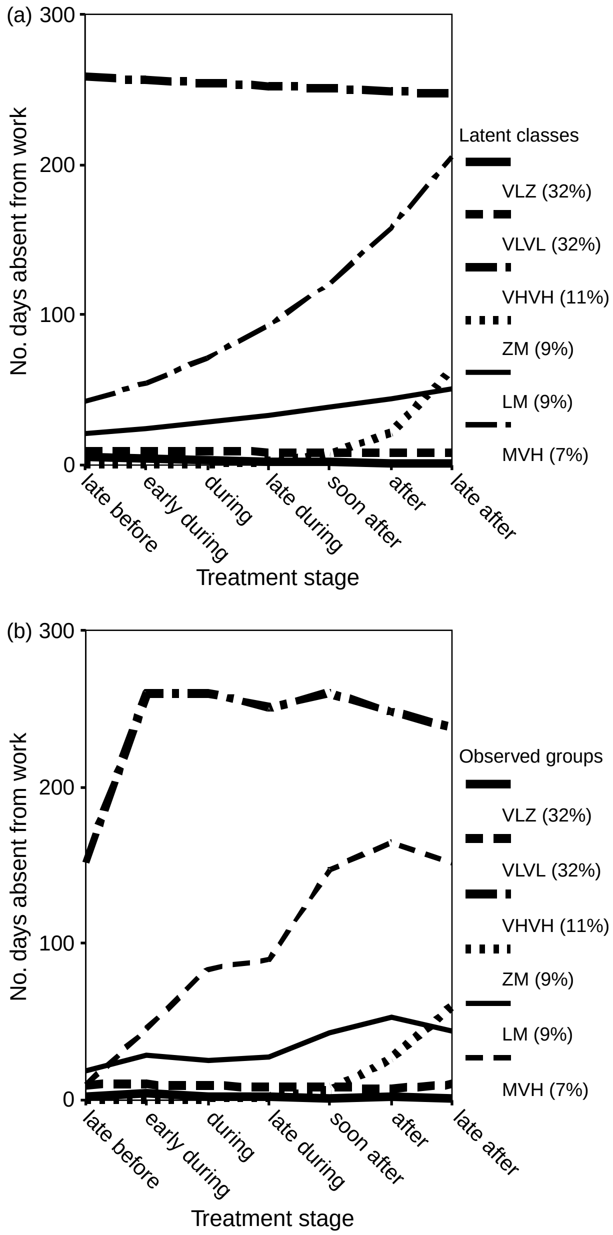

Figure 6. Modeled trajectories for number of days absent from work for the six latent classes (upper panel) and observed trajectories for the corresponding six groups of patients (lower panel). (H = high; L = low; M = medium; VH = very high; VL = very low; Z = [close to] zero).

in decreasing order of association strength: VIS, GAF-C, GAF-L, Axis I diagnosis, and treatment type (all ps $<.01$). Figure 7 shows how these variables combined to predict class membership. Four covariate segments were obtained. Segment A consisted of 65 patients (30%) with no persons from class VHVH and overrepresentation of class VLZ members and slight overrepresentation of class LM: These were persons with no to moderate vocational impairment and no Axis I diagnosis and with psychotherapy as their treatment. In contrast, segment B (11%) had an overrepresentation of classes VLVL and ZM and no persons from classes VHVH, LM, or MVH: These were patients in psychoanalysis without any Axis I diagnosis and with no impairment vocationally. Segment C was the largest subgroup (111 persons; 51%): low to moderate vocational impairment yet with an Axis I diagnosis. There was no particular class with clear over- or underrepresentation. Segment D, finally, was a quite small group (16 persons; 7%) and consisted mainly of individuals from classes VHVH, LM, and MVH, and the segment was distinguished by their high scores on the VIS, from severe to complete vocational impairment.

Overlap Among the Clusters

A clinically significant question is, To what extent do the clusters on the six HCUVs overlap? Are we observing basically the same clusters again and again, or were the heavy medicine consumers, for instance, other people generally than those who have frequent days off from work or those who are frequent inpatients? Is there a strong general HCU factor that might account well enough for the six clusterings when run into an LCA? A principal-components transformation of the HCUVs indicated that two components accounted each for more than any single variable (eigenvalues >1). In view of this, it is indeed reasonable to expect that there was some overlap among the clusters. On the other hand, the two components together accounted for only little more than 50% of the total variance, so there is unique variance enough to expect much less than perfect overlap. Confronted with the hazards of applying parametric analyses to the HCUVs except for exploratory purposes, we chose to proceed along another route.

A series of cross-table analyses showed that the significant specific co-occurrences between the classifications were infrequent and rather complicated. As one example, when we cross-tabulated the six clusters based on somatic consultations with the six clusters based on absence days, there were no more than three cells among the 36 with more than the expected frequencies (standardized adjusted residuals >2). One such case was between clusters with similar developments (consultations VLZ and absence VLZ), but the two others were inconsistent, between a cluster with reduction in consultations (classes MM or VHH) and one with increasing days of absence (LM).

Then, as a kind of omnibus cross-table analysis, we performed an LC nonparametric cluster analysis with cluster membership on the six HCUVs as nominal clustering variables and with the same

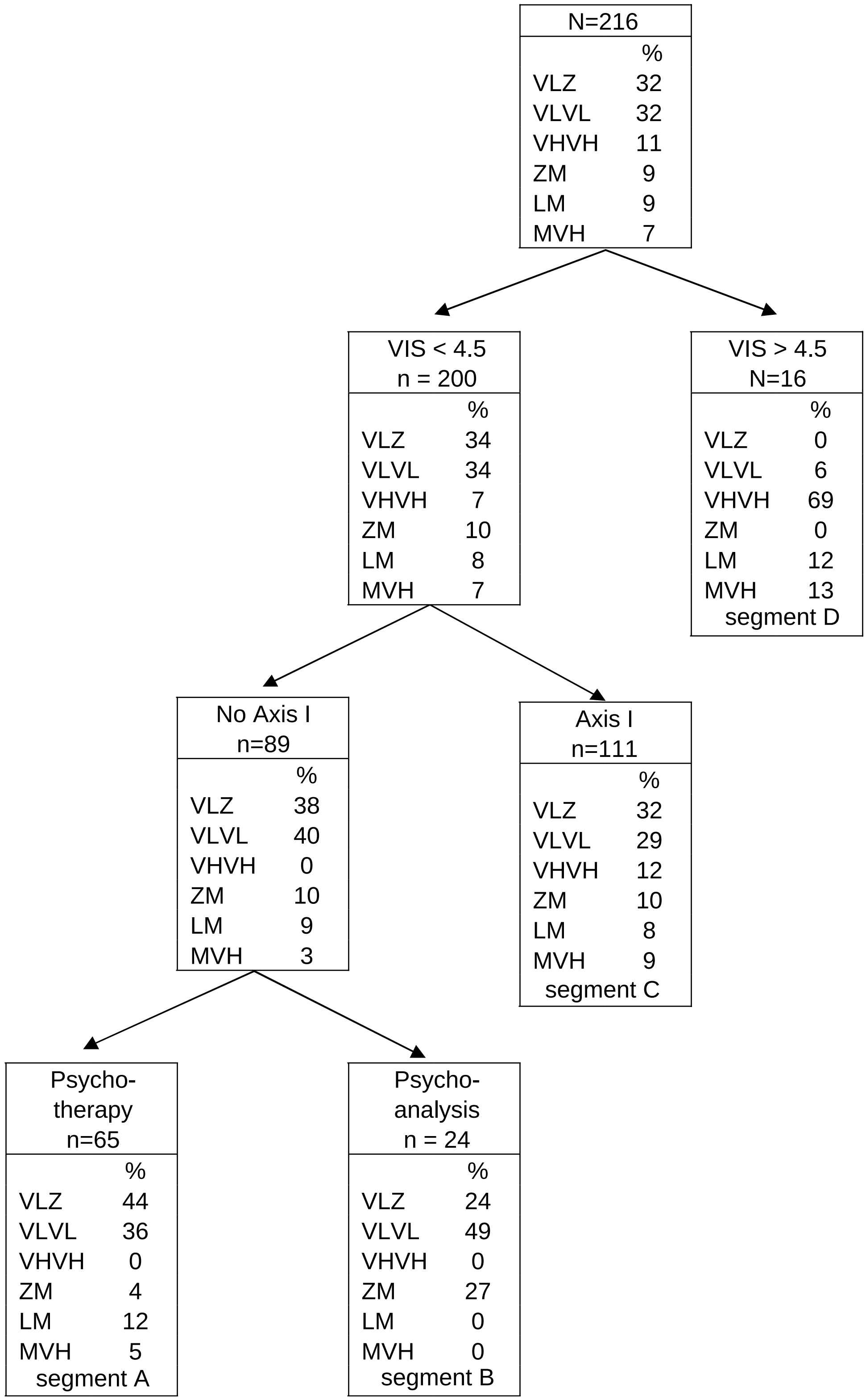

Figure 7. SI-CHAID-generated tree diagram relating the six latent absent days classes to the nonredundant covariates. (# = class number; VIS = Vocational Impairment Scale. Specific segments refer to groups of combinations of covariates; H = high; L = low; M = medium; VH = very high; VL = very low; Z = [close to] zero).

inactive covariates as used in the LC regression analyses. To include absence days among the HCUVs, this analysis was run on the subsample of 216 persons belonging to the labor force. The BIC suggested a two-class solution. These two second-order clusters accounted for not more than between

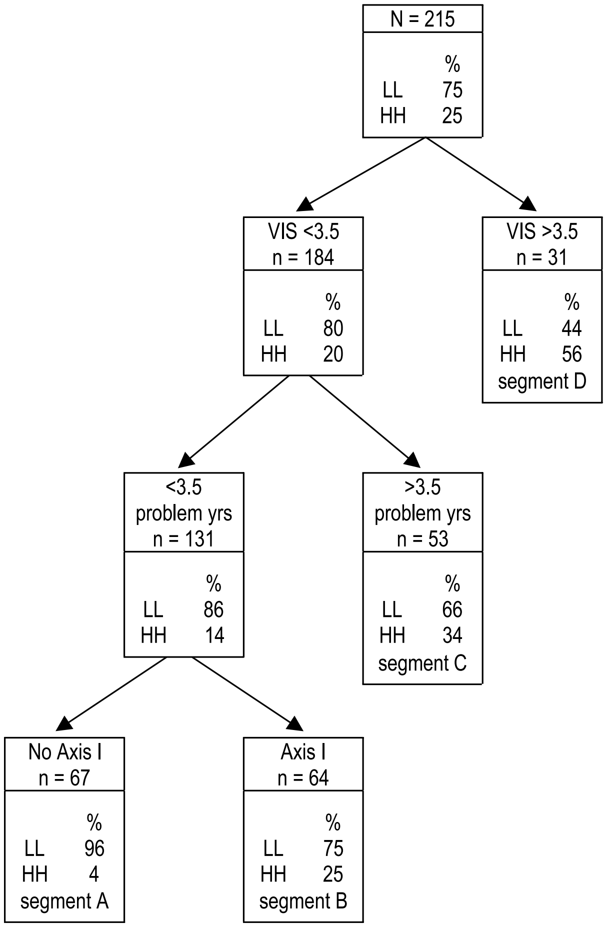

Figure 8. SI-CHAID-generated tree diagram relating the two latent health care utilization clusters to the non-redundant covariates. VIS = Vocational Impairment Scale; Axis I = DSM-IV Axis I diagnosis. Specific segments refer to groups of combinations of covariates.

2% (somatic consultations clusters) and 25% (psychiatric consultations clusters; *Mdn* = 14%) of the variance in the six first-order clusterings.[7] Solutions with progressively higher numbers of second-order clusters until nine (when the number of degrees of freedom was 0) revealed that not more than 24% of the variance was explained among the somatic consultations clusters, which was the least again, and not more than 42% of the variance among the psychiatric consultations clusters, which was the most again (*Mdn* = 31%). We concluded that the first-order clusterings, except for medicine consumption, had too much unique variance to be well accounted for by any second-order clustering.

Yet, for the sake of exploration, we did pursue the two-cluster solution. The classes were easily identifiable as one (LL; 75%) with light or very light consumers of health care (significant overrepresentation of ZZ on psychiatric consultations, LL and MM on medicine consumption, VLVL on absence) and one (HH; 75%) with persons with heavier consumption (significant overrepresentation of VHVH and HH on medicine consumption). The significant overrepresentations in one cluster were balanced by significant underrepresentations of the other cluster. Cluster membership was significantly ($p < .01$) associated with VIS, GAF-L, GAF-C, Axis I diagnosis, and number of problem years. Given the associations among these covariates, low VIS scores, few problem years, and no Axis I diagnosis formed a covariate segment (A) with 67 persons (31%), of whom 96% belonged to cluster #1. The proportion of cluster #1 persons shrunk to 75% among the 64 (30%) persons with low VIS, few problem years but at least one Axis I diagnosis (segment B); to 66% as the number of problem years increased to four or more (irrespective of diagnosis) (segment C) and to 44% if the person had a pretreatment VIS score > 3 (severe to complete impairment), irrespective of problem years or diagnosis (segment D). The tree diagram is given in Figure 8.

Discussion

The results support the proposition that change in HCU during psychotherapy is heterogeneous. Patients differ systematically not only in their levels of HCU but also in their change patterns during psychotherapy. For almost each of the HCUVs, there are classes with upward trends as well as classes with downward trends and classes with little or no change at high or low levels of HCU. Given this heterogeneity, the main effect is a severe misrepresentation, representative for few, if any, of the patients. If we take consultations with MDs for somatic ailments as an example, changes in none of the five classes coincided with the change in the undivided sample, although, of course, the influence of the classes on the sample mean reflects their relative sizes.

Also in line with our general hypothesis, the heterogeneity may partly be predicted on the basis of patients' pretreatment characteristics. These characteristics were in most cases related to their mental condition and not as much to demographic and other kinds of sociological variables. Although some of the interactions with the VIS appear tautological (e.g., inability to work is tantamount to heavy absenteeism, although it should be noted that the VIS was pretreatment), some of the other interactions suggest that there are critical personality variables involved. In the absence of any personality assessment, we do not know which these may be. Because the LC variable acts as a moderator variable (Magidson & Vermunt, 2002), these interactions should be considered as aptitude × outcome effects (Blatt & Felsen, 1993).

Also, in a qualitative sense, the changes during psychotherapy were heterogeneous. HCU is multidimensional, and different classes of patients display different patterns of utilization and change. Although the associations among the HCUVs suggested two components, there was too much uniqueness to the different HCUV clusterings for any one to account for much of any other. Therefore, the conventional focus of cost offset studies on inpatient care and absenteeism (Gabbard et al., 1997; Mumford et al., 1984) may not generate a representative picture of the human implications in all their aspects. From an economic point of view, outpatient consultations and medicine consumption may be less important, but there are certainly other important, human consequences than cost offset. As Lazar et al. (2006) noted, outpatient care does not affect costs as dramatically as inpatient care but, until it becomes very bad, ill health has a more immediate effect on outpatient care than on inpatient care.

This qualitative heterogeneity only partly conforms to the suggestion of Chiles et al. (1999) that the effects of psychotherapy on HCU depend on whether such effects were its primary purpose or not. If we assume that changes in mental HCU are closer to the aims of psychotherapy than changes in somatic HCU, we should expect more cases with decreasing patterns with the former than with the latter. There were no clear indications to that effect when outpatient consultations were concerned but an obvious difference with inpatient weeks.

The majority of patients were people with initially very low levels of HCU, generally, without much room for improvement. These patients had, nevertheless, high levels of psychological distress; more than 80% of the total sample was initially below the normal mean in self-rated health. Paradoxically, their problem from a health care view was sickness presenteeism than absenteeism (i.e., going to work despite the fact that one's current state of health warrants sick leave; Aronsson & Gustafsson, 2002; Aronsson, Gustafsson, & Dallner, 2000). In psychotherapy populations where these people are frequent, using HCUVs as measures of outcome may be irrelevant and indeed misleading, particularly when decreasing levels of HCU are expected. Even, in these cases, a positive effect of psychotherapy might rather be an increase. In fact, Lazar et al. (2006) found an upward trend for patients with very low initial levels of absenteeism and a downward trend for patients with initially higher levels, the two groups converging towards the "normal", national capacity rate. Regression to the mean could be ruled out in that design (Blomberg et al., 2001).

We have to ask whether the outcome heterogeneity among patients is not merely error variance. First, the fact that the class variable in most of our analyses was associated with one or more covariates contradicts this. This could not have happened (except by chance, of course) if one (or both) of the variables had very low reliability. Second, the proportion of true variance in a test variable is estimated by the reliability of the test, and this applies as well to the within-groups variance in a treatment study (Lyons & Howard, 1991). Lazar et al. (2006) estimated lower bound reliabilities as the between-consecutive-years correlations. Keeping in mind that these are stability coefficients rather than reliabilities, the correlations were generally satisfactory, from moderate to high in size. Inpatient weeks in somatic care had very low stability coefficients, however. We do not regard these as low reliabilities but rather as indications of the fact that few people in our sample were repeatedly hospitalized.

Third, there is the validity of these self-reports, which of course reflects their reliability. Lazar et al. (2006) compared self-reports with official records for the same time periods and concluded that the agreement was in some cases impressive, considering the limitations of the official records. We conclude, therefore, that there was sufficient systematic or reliable heterogeneity in the outcome trajectories of these outcome variables.

An interesting question is whether the heterogeneity found is actually due to variations among the patients or their therapists. Therapists would probably like to believe that the most potent source of variation is the patient, and that is also the general opinion among psychotherapy researchers (Norcross & Lambert, 2006). However, evidence is accumulating for the variability among therapists as well (Beutler, 1997; Blatt, Sanislow, Zuroff, & Pilkonis, 1996; Crits-Christoph et al., 1991; Crits-Christoph & Mintz, 1991; Kim, Wampold, & Bolt, 2006; Lafferty, Beutler, & Crago, 1989; Lambert, 1990; Luborsky et al., 1986; Luborsky, McLellan, Diguer, Woody, & Seligman, 1997; Luborsky, McLellan, Woody, O'Brien, & Auerbach, 1985; Wampold & Brown, 2005). Also, two studies using LC regression analysis reported large between-therapists differences (Sandell et al., 2006a, 2006b). Obviously, in this study, as for most psychotherapy studies, there was a partial confounding between patients and therapists, because each patient had only one therapist, although each therapist may have had more than one patient, and there was no random selection between the two. Although our design is not ideal, we tried to approach the problem by comparing the across-classes variances of change slopes for patient classes with therapist classes for some of the variables. When medical consultations were concerned, the variance across the six patient classes was greater than across six therapist classes (0.52 vs. 0.38), with absence days essentially equal (0.47 vs. 0.50) and with psychiatric consultations substantially less (0.21 vs. 0.41). Thus, it seems that the responsibilities or contributions of the two partners may vary with the kind of outcome. More research is certainly needed on this important issue.

What if psychotherapy outcomes in general are heterogeneous like these, whether this depends on the therapists or the patients? We submit that, to the extent that there are systematically different outcome subgroups, main effects analyses do, in fact, distort what has happened in a treatment. Concluding that mean change was, for example, 0 is indeed very different from the conclusion that 50% of the patients had -1 unit of change and the rest had 1, especially if these subgroups were associated with person characteristics.

In the face of heterogeneous outcomes like the ones reported here, how may we formulate the results and conclusions of outcome studies? The mean change represents the expected value but will become more useless the greater the heterogeneity. The more reasonable approach, it seems to us, is to accept that sample outcome is complex and present a decomposition of the sample using some variant of class analysis. If one has to simplify complex matters, one may report the proportion of patients in the treatment who displayed positive, or nonnegative, change. In this study, this varied between the outcome variables, from 100% in psychiatric hospitalization to 69% in somatic hospitalization when classes with nonnegative trends were collapsed and counted. Besides 75% for absenteeism, a typical number was 85% to 86% for somatic and psychiatric consultations and medicine consumption.

A reasonable question is then, To what extent are the developments in the negative cases negative effects or unwanted side effects of the therapies? Although there certainly may be single such cases, a review of the patients' comments has not led us to believe that that is a frequent phenomenon. Although the psychosomatic perspective probably should be the default in the psychotherapy context, it is a fact of life that people fall sick for very different reasons, of which not all can be related to the patient's lifestyle, ways of coping, temperament, and so on.

Cronbach (1957) made the distinction between two subdisciplines of scientific psychology: experimental and correlational psychology. A related distinction can be made in terms of two other perspectives: those of the general and the differential psychologies. General psychology focuses on the average statistic in search of that which is general among humans, regarding the variation around the average primarily as a nuisance and using it as the basis for error estimates. The classical learning curve is an example. Differential psychology has made this variation its very subject matter in its description of ways in which human beings are different and attempts to explain this, as in research on intelligence and personality factors. More or less unwittingly, psychotherapy outcome research has adopted the perspective of general psychology. Although it is difficult for some of us to apply two perspectives at the same time, the present findings suggest that one should at least try.

Notes

[1] The too-small low-dose group was excluded to simplify the analysis of treatment type to a binary variable.

[2] The model for counts is a log-linear Poisson regression, $\ln(\mu)_{ti} = \beta_0 + \beta_1 {}^\star t_i$ or
$(\mu)\mathrm{ti} = \exp(\beta0 + \beta1 \star \mathrm{ti})$
where μ is the population mean, β_0 and β_1 are intercept and slope estimates, respectively, and t_i is the ordinal stage number. Thus, the natural logarithm of μ changes linearly with β_1 per stage, whereas μ will change nonlinearly in most cases. Because the estimates may be difficult to comprehend at first glance, we also report the accumulated change in raw scores from the year before treatment to the third year after termination and, in cases where the trend is reasonably linear, the average change per stage.

[3] The change pattern of each cluster is characterized in terms of the approximate relative levels, initially and finally, using the following descriptors: H = high; L = low; M = medium; VH = very high; VL = very low; Z = (close to) zero.

[4] The covariates are ordered according to the strength of their association with the classification; the covariate with the strongest association is given first.

[5] The paradoxical fact that the entire group developed in the increasing direction but all four latent classes developed in the decreasing direction deserves some reflection. The two-, three-, and five-class solutions—but not the four-class solution—had one class with increasing number of weeks. This is not likely a software artifact because another software yielded exactly the same pattern. Possibly patients with increasing number of weeks are divided among the four classes without substantially affecting the class average, although they contribute to increased within-class heterogeneity/error.

[6] For category scales, the adjacent category logit model is specified. The β_1 is a log-odds ratio, meaning that the odds of category j versus category $j - 1$ ($j = 1, 2 \ldots 5$ for a five-step scale) changes by $\exp(\beta_1)$ per stage. The odds are assumed constant for all pairs of j and $j - 1$.

[7] For a nominal variable, the variance explained represents the weighted average of separate squared multiple correlations for each category (class) taken as a dichotomous variable.

References

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Antonovsky, A. (1987). *Unraveling the mystery of health: How people manage stress and stay well.* San Francisco: Jossey-Bass.

Aronsson, G., & Gustafsson, K. (2002). Sickness presenteeism—Prevalence and trends. *Arbete och Hälsa, 8.* Solna, Sweden: Arbetslivsinstitutet.

Aronsson, G., Gustafsson, K., & Dallner, M. (2000). Sick but yet at work. An empirical study of sickness presenteeism. *Journal of Epidemiology and Community Health*, *54*, 502–509.

Beutel, M. E., Rasting, M., Stuhr, U., Rüger, B., & Leuzinger-Bohleber, M. (2004). Assessing the impact of psychoanalyses and long-term psychoanalytic therapies on health care utilization and costs. *Psychotherapy Research*, *14*, 146–160.

Beutler, L. E. (1997). The psychotherapist as a neglected variable in psychotherapy: An illustration by reference to the role of therapist experience and training. *Clinical Psychology: Science and Practice*, *4*, 44–52.

Blatt, S. J., & Felsen, I. (1993). Different kinds of folks may need different kinds of strokes: The effect of patients' characteristics on therapeutic process and outcome. *Psychotherapy Research*, *3*, 245–259.

Blatt, S. J., Sanislow, C. A., Zuroff, D. C., & Pilkonis, P. A. (1996). Characteristics of effective therapists: Further analyses of data from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, *64*, 1276–1284.

Blomberg, J., Lazar, A., & Sandell, R. (2001). Long-term outcome of long-term psychoanalytically oriented therapies: First findings of the Stockholm Outcome of Psychotherapy and Psychoanalysis Study. *Psychotherapy Research*, *11*, 361–382.

Bouwmeester, S., Sijtsma, K., & Vermunt, J. K. (2004). Latent class regression analysis for describing cognitive developmental phenomena: An application to transitive reasoning. *European Journal of Developmental Psychology*, *1*, 67–86.

Breyer, F., Heinzel, R., & Klein, T. (1997). Kosten und Nutzen ambulanter Psychoanalyse in Deutschland [Cost and benefit of outpatient psychoanalysis in Germany]. *Zeitschrift für Gesundheitsökonomie und Qualitätsmanagement*, *2*, 59–73.

Chiles, J. A., Lambert, M. J., & Hatch, A. L. (1999). The impact of psychological interventions on medical offset: A meta-analytic review. *Clinical Psychology: Science and Practice*, *6*, 204–220.

Clarkin, J. F., & Levy, K. N. (2004). The influence of client variables on psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 194–226). New York: Wiley.

Crits-Christoph, P., Carroll, K., Perry, K., Luborsky, L., McLellan, A. T., Woody, G. E., et al. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research*, *1*, 81–91.

Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology*, *59*, 20–26.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684.

Derogatis, L. R., & Lazarus, L. (1994). SCL-90-R, Brief Symptom Inventory, and matching clinical ratings scales. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 217–248). Hillsdale, NJ: Erlbaum.

Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science*, *19*, 1–15.

Dührssen, A. (1962). Katamnestische Ergebnisse bei 1004 Patienten nach analytischer Psychotherapie [Follow-up results among 1004 patients after analytical psychotherapy]. *Zeitschrift für Psychosomatische Medizin*, *8*, 94–113.

Dührssen, A., & Jorswieck, E. (1965). Eine empirisch-statistische Untersuchung zur Leistungsfähigkeit psychoanalytischer Behandlung [An empirical-statistical study of the potential effects of psychoanalytical treatment]. *Nervenarzt*, *36*, 166–169.

Follette, W., & Cummings, N. A. (1967). Psychiatric services and medical utilization in a prepaid health plan setting. *Medical Care*, *1*, 25–35.

Gabbard, G. O., Lazar, S. G., Hornberger, J., & Spiegel, D. (1997). The economic impact of psychotherapy: A review. *American Journal of Psychiatry*, *154*, 147–155.

Keller, W., Westhoff, G., Dilg, R., Rohner, R., & Studt, H. H. (2001). Wirksamkeit und Kosten-Nutzen-Aspekte ambulanter (jungianischer) Psychoanalysen und Psychotherapien—Eine katamnestische Studie [Effectiveness and cost-benefit aspects of outpatient (Jungian) psychoanalyses and psychotherapies—A follow-up study]. In U. Stuhr, M. Leuzinger-Bohleber & M. E. Beutel (Eds), *Langzeitpsychotherapie—Perspektiven für Therapeuten und Wissenschaftler* (pp. 343–355). Stuttgart, Germany: Kohlhammer.

Kim, D.-M., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research*, *16*, 144–160.

Lafferty, P., Beutler, L. E., & Crago, M. (1989). Differences between more and less effective psychotherapists: A study of select therapist variables. *Journal of Consulting and Clinical Psychology*, *57*, 76–80.

Lambert, M. J. (1990). The individual therapist's contribution to psychotherapy process and outcome. *Clinical Psychology Review*, *9*, 469–485.

Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists. In J. C. Norcross & M. R. Goldfried (Eds), *Handbook of psychotherapy integration* (pp. 94–129). New York: Basic Books.

Lazar, A., Sandell, R., & Grant, J. (2006). Do psychoanalytic treatments have positive effects on health and health care utilization? Further findings of the Stockholm Outcome of Psychotherapy and Psychoanalysis Project (STOPPP). *Psychotherapy Research*, *16*, 51–66.

Lazar, A., Sandell, R., & Grant, J. (in press). Subjective health and ill health-related behaviour. *Psychology and Psychotherapy: Theory, Research and Practice*.

Luborsky, L., Crits-Christoph, P., McLellan, A. T., Woody, G., Piper, W., Liberman, B., et al. (1986). Do therapists vary much in their success? Findings from four outcome studies. *American Journal of Orthopsychiatry*, *56*, 501–521.

Luborsky, L., McLellan, A. T., Diguer, L., Woody, G., & Seligman, D. A. (1997). The psychotherapist matters: Comparison of outcomes across twenty-two therapists and seven patient samples. *Clinical Psychology: Science & Practice*, *4*, 53–65.

Luborsky, L., McLellan, A. T., Woody, G. E., O'Brien, C. P., & Auerbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry*, *42*, 602–611.

Lyons, J. S., & Howard, K. I. (1991). Main effects analysis in clinical research: Statistical guidelines for disaggregating treatment groups. *Journal of Consulting and Clinical Psychology*, *59*, 745–748.

Magidson, J. (2005). *SI-CHAID 4.0 user's guide*. Belmont, MA: Statistical Innovations Inc.

Magidson, J., & Vermunt, J. K. (2002). *Latent class models*. Belmont, MA: Statistical Innovations Inc. Retrieved from www.statisticalinnovations.com

Mumford, E., Schlesinger, H. J., Glass, G. V., Patrick, C., & Cuerdon, T. (1984). A new look at evidence about reduced cost of medical utilization following mental health care treatment. *American Journal of Psychiatry*, *141*, 1145–1158.

Norcross, J. C., & Lambert, M. J. (2006). The therapy relationship. In J. C. Norcross, L. E. Beutler & R. F. Levant (Eds), *Evidence-based practices in mental health. Debate and dialogue on the fundamental issues* (pp. 208–218). Washington, DC: American Psychological Association.

Sandell, R., Blomberg, J., Lazar, A., Carlsson, J., Broberg, J., & Schubert, J. (2000). Varieties of long-term outcome among patients in psychoanalysis and long-term psychotherapy. A review of findings in the Stockholm Outcome of Psychoanalysis and Psychotherapy Project (STOPPP). *International Journal of Psychoanalysis*, *81*, 921–942.

Sandell, R., Carlsson, J., Schubert, J., Broberg, J., Lazar, A., & Grant, J. (2006a). Therapist attitudes and patient outcomes: III. A latent class (LC) analysis of therapists. *Psychology and Psychotherapy. Theory, Research, and Practice*, *79*, 629–647.

Sandell, R., Carlsson, J., Schubert, J., Broberg, J., Lazar, A., & Grant, J. (2006b). Therapists' therapies. The relation between therapists' personal/training therapies and patient outcome in long-term psychotherapy and psychoanalysis. *Psychotherapy Research*, *16*, 306–316.

Wampold, B. E. (2001). *The great psychotherapy debate. Models, methods, and findings*. Mahwah, NJ: Erlbaum.

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology*, *73*, 914–923.

Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry*, *33*, 1111–1115.

Weissman, M. M., Prusoff, B. A., Thompson, W. D., Harding, P. S., & Myers, J. K. (1978). Social adjustment by self-report in a community sample and in psychiatric outpatients. *Journal of Nervous and Mental Disease*, *166*, 317–326.

Vermunt, J. K., & Magidson, J. (2005). *Latent GOLD 4.0 user's guide*. Belmont, MA: Statistical Innovations Inc.

Vermunt, J. K., & van Dijk, L. A. (2001). A non-parametric random coefficient approach: The latent class regression model. *Multilevel Modelling Newsletter*, *13*, 6–13.