Classification and Regression Tree Uncovered Hierarchy of Psychosocial Determinants Underlying Quality Of Life Response Shift in HIV/AIDS

Yuelin Li, PhD.[1] and **Bruce Rapkin, PhD.**[2]

[1]Department of Psychiatry and Behavioral Sciences, Memorial Sloan-Kettering Cancer Center. 641 Lexington Ave. 7th Floor. New York, NY 10022, Tel: 646-888-0047, Fax: 212-888-2959, liy12@mskcc.org

[2]Department of Epidemiology and Population Health, Albert Einstein College of Medicine (AECOM) of Yeshiva University, Bronx, NY. brapkin@aecom.yu.edu

Abstract

Objectives—Rapkin and Schwartz [1] define response shift as otherwise unexplained, discrepant change in HRQOL that is associated with change in cognitive appraisal. In this paper, we demonstrate how a Recursive Partitioning (RPART) regression tree analytic approach may be used to explore cognitive changes to gain additional insight into response shift phenomena.

Study Design and Setting—Data are from the "Choices in Care Study," [2], an evaluation of HIV+ Medicaid recipients' experiences and outcomes in care (N = 394). Cognitive assessment was based on the Quality of Life Appraisal Battery. HRQOL was measured by the MOS-SF36v2 [3].

Results—We used RPART to examine six-month change in MOS mental composite as a function of changes in appraisal, after controlling for patient characteristics, health changes and intervening events. RPART identified nine distinct patterns of cognitive change, including three associated with negative discrepancies, four with positive discrepancies, and two with no discrepancies.

Conclusion—RPART classification provides a nuanced treatment of response shift. This methodology has implications for evaluating programs, guiding decisions and targeting care.

Keywords

response shift; health related quality of life; classification and regression trees

1. Introduction

Converging evidence shows that response shift can strongly affect how in individuals appraise their Health-Related Quality Of Life (HRQOL) [1,4,5,6]. Response shift typically appears in counterintuitive findings—individuals with severe chronic illnesses reporting equal or better HRQOL scores than healthy individuals or individuals with less severe illness (e.g., [7,8]). For example, the general public assigned a 0.39 HRQOL to dialysis whereas dialysis patients assigned their own HRQOL at 0.56 (on a 0 – 1 scale where 0 represents death and 1 represents

perfect health) [9]. This and similar paradoxical findings bring into question what HRQOL assessments are really measuring. Measurement imprecision and response bias do not fully explain the phenomena [1]. The theory of response shift posits that it constitutes a change in the meaning of one's self-evaluation of the QOL construct due to recalibration, repriortization, and/or reconceptualization [6,10]. These constructs are related to work on idiographic quality of life assessment [4,11,12,13]. The theoretical and measurement foundations of these constructs are well documented [5,6,14].

Rapkin and Schwartz [1] describe the assessment strategies to probe respondents on their evaluation of the meaning of QOL. They propose operationalizing response shift as change in HRQOL that cannot be explained by changes in overt health status, resources or life events, but that can be associated with change in cognitive appraisal. Their data-analytic strategy was based primarily on linear regression to estimate the extent to which residual QOL changes are associated with appraisal change [4]. However, there is no intrinsic reason that these relationships must be linear. Rapkin and Schwartz' notion of a final "combinatorial algorithm" that people use to summarize their experiences into HRQOL ratings explicitly posits complex interactions among constituents of appraisal. For example, an individual may report better quality of life than expected given their health status by ignoring problems, emphasizing positive experiences, selecting favorable targets for self-comparison, and/or focusing on less ambitious goals. Each of these processes may operate alone or in combination to represent distinct types of response shift.

There are obvious drawbacks to using linear regression to examine relationships involving appraisal processes that are intrinsically non-linear. Classification and Regression Trees (CART) methods are a suitable alternative to linear regression in elucidating potentially complex interactions [15]. Using an iterative algorithm, respondents are classified into increasingly homogeneous subgroups with similar changes in cognitive appraisal profiles, allowing a more nuanced interpretation of how cognitive appraisal can influence HRQOL. The broad goal of this article is to demonstrate an empirical technique to identify prevalent patterns of cognitive changes that can account for residual variance in HRQOL change scores. Patterns of appraisal identified in this way represent different manifestations of response shift.

2. Methods

2.1 HIV/AIDS Choices in Care Study

The study was developed by investigators at our respective institutions in conjunction with the New York State Department of Health AIDS Institute to evaluate the impact of the HIV Special Needs Plans, as part of an evaluation of patient-reported outcomes and experiences in care reported by HIV+ Medicaid recipients in New York State. Detailed data collection plans are summarized elsewhere [2,16]. Institutional Review Boards approved the study.

Interviews were conducted in either Spanish or English, in person or by telephone, according to patient preference. The primary HRQOL assessment was the SF36.v2 [3], assessed at baseline (approximately six weeks post enrollment), and at 6 and 12 months post baseline. Changes in cognitive appraisal processes were assessed at these time points using the Quality of Life Appraisal Battery [1]. The baseline interview also included measures of demographics, behavioral risks and health history.

2.2 Quality of Life appraisal battery

Following Rapkin and Schwartz [1], the Quality of Life Appraisal Battery included four components: 1) persons' *frame of reference* for considering HRQOL as assessed by six probes designed to tap different motivational themes, including achievement, maintenance,

prevention, problem solving, disengagement, and acceptance. For example, respondents were asked about "the main things they want to accomplish", "problems they want to solve", and "things they are trying learn to accept," in order to have their best possible quality of life. Verbatim responses to these probes, or "goal statements," were coded and analyzed to extract "goal attributes" (described below). Additionally, we assessed 2) how persons *sample experiences* within that frame, assessed by 13 items on, e.g., whether or not the persons evaluated HRQOL by "thinking about the worst possible moments" within that frame; 3) how persons evaluate experiences using different *standards of comparison* by 9 items on, e.g., whether they compared themselves with "other people living with HIV"; and 4) how persons summarize and combine evaluations to describe HRQOL by using a combinatory algorithm of 16 items on, for example, whether they were thinking about "how well you've been doing, how hard it has been, both or neither?".

2.3 Coding and summarizing goal statements to assess frame of reference

From the open-ended assessments of frame of reference we collected over 6,700 goal statements at baseline and 6 months (plus an additional 1,458 from our first wave of 12 months follow-up interviews, which were coded in this group, but not reported here). Content analysis of these responses was accomplished through a two-stage process which is briefly summarized below. Complete documentation of goal coding, Kappa reliability and components analysis is available from the authors upon request.

In the first step, we selected at random just over 1/3 of all responses (2,638). Each selected goal was given to two of 13 judges (students and faculty in our department), following an allocation scheme to ensure that an equal number of overlapping goals assigned to each pair of judges. Each judge independently sorted about 405 goal statements into homogeneous categories, with the sole criterion being that statements within a category must be "similar in all important ways". Judges then recorded the "goal attributes" that they used to make distinctions among categories, including life domains, motivations, and health -relevance. After completion of independent sorting, all judges met to compare their derived dimensions. In general, there was strong agreement in the major distinctions among life domains and in prevalent fine-grain distinctions. Judges primarily differed in how specific to be in certain sub-domains (for example, to distinguish concerns about specific family members from those pertaining to the family in general). Based on this discussion, we derived a consensus set of 24 binary goal attributes. All goal statements could be characterized using combinations of these goal attributes. We calculated kappa for each of the 24 codes, to determine whether or not pairs of judges agreed on the presence or absence of each goal attribute in their initial sort of goal statements. Collapsing across dyads, we found that 11 of 24 categories exceed kappa = .70, another 4 exceed kappa =.59, 6 exceeded kappa = .35, and 3 codes (representing only 3.76% of coded statements) did not differ from chance.

Following derivation of goal attributes, the remaining 2/3 (5,148) goal statements were assigned to eleven judges. We assigned a random 20% of these goal statements (1,030) as a reliability sample, allocated evenly to all possible pairs of judges. Reliability coefficients for 13 of 24 categories exceeded kappa = .70, another 6 exceed kappa =.50, 3 exceeded an acceptable level of kappa = .39, and 2 categories (representing only 1.19% of coded statements) were not different from chance. Based on these results, final goal attributes were coded for each goal statement. Note that in final coding, we resolved disagreements among judges by assuming that differences were due to errors of omission (one judge indicated a code that the other did not).

Our next step involved combining scores across all of individual's goal statements at baseline and separately at six-months, to characterize current priorities and concerns at each time of measurement. Our goal at this step was to achieve a parsimonious data reduction while retaining

as much information as our data would permit. Our coding system yielded a binary vector describing the presence or absence of 24 different goal attributes for each goal statement. Recall that goal statements were elicited by six different motivational probes. Thus, for the 9 most prevalent codes, we calculated subtotals representing the occurrence of each goal attribute for statements elicited by each motivational theme. For example, this cross-classification allowed us to distinguish among goals about solving money problems, earning more money or learning to live more frugally. The 54 variables formed by the cells of this cross-classification of 9 major goal attribute codes by 6 motivational themes fully accounted for 74% of all responses. These represented our primary goal attributes. We reduced these 54 variables by conducting a two-stage principal components analysis, first summarizing endorsement rates of codes within each of the six motivational themes (retaining 57% to 89% of total variance in each set), and then combining the 32 first-order components from these six analyses in a single second-order principal components analysis (retaining 60% of variance among the first-order components). Second-order analysis yielded sixteen major goal attribute factors. These components are listed in Table 2.

The remaining 15 goal attribute codes were less prevalent, so we simply tallied the total number of times content codes occurred for each individual at a given time of measurement, without subtotaling by eliciting theme. Principal components analysis yielded seven relatively independent components after promax rotation, summarizing 58% of the variance among these 15 codes. These seven subsidiary goal attribute dimensions are also listed in Table 2. Substantively, we think of the primary goal content factors as capturing the individuals' status in broadly shared areas of concern, while the subsidiary dimensions reflect more particular concerns that may nonetheless have an important influence on individuals' appraisal of quality of life.

2.4 Scoring other domains of quality of life appraisal battery

The other three parameters of quality of life appraisal were analyzed by a series of principal component analyses to map the items of *sample experiences* to 5 factors, *standards of comparison* to 3 factors, and *combinatory algorithms* to 7 factors. Generally, principal components with eigenvalues greater than or equal to 1 were retained. Table 3, Table 4, and Table 5 summarize the total variance accounted for by the retained eigenvalues, and the rotated factor loadings of the items. Standardized factor scores were calculated and entered into the analysis. Take the *combinatory algorithms* scale in Table 5 as an example, respondents were prompted "When you answered today, did you think more about …" and they rated the extent to which they thought about "Things that are disappointing to you", "How hard it has been", "Things that make you feel worried", and so on. These three items had high factor loadings on the first factor, which was thus labeled as "Negative Experiences, Feelings & Worries".

2.5 Changes in quality of life appraisal

Because the QOL appraisal subscales were standardized, all sub-domains were thus mapped onto a comparable scale of mean zero and unit standard deviation. A respondent with a zero "reacting to recent flare-ups" score, for example, represents an appraisal through recent disease flare-ups at the sample average. Changes in QOL appraisal were thus operationalized as changes in the standardized scores. In principle, the changes in standardized scores can be thought of as changes in effect size units [17,18], thereby simplifying comparisons made across multiple QOL appraisal domains on arbitrary raw scales. We felt that it would facilitate interpreting the changes in appraisal by considering a set of crude by practical cut-offs. We considered a 0.75 standard deviation change a 'moderate' change in appraisal, a 1.0 change a 'moderately large' change, and a 1.5+ change a 'large' change. The 'large' change is conveniently twice as large as the 'moderate' change. These cut-offs are more conservative

than the conventional effect size indexes [18] (e.g., 0.80 as a 'large' effect), and we believe that they help track the numerous and complex pattern of splits in the rpart analysis.

2.6 HRQOL discrepancy score

Following Rapkin and Schwartz' [1] formulation, in order to examine HRQOL response shift, it was first necessary to derive a discrepancy score in the HRQOL, to determine how much the observed score differed from an expected value. Response shift arises when the observed changes in HRQOL scores deviate systematically from the expected HRQOL changes due to health-related events. A simplistic example illustrates the basic conceptual premises. If a person experiences more symptoms related to HIV/AIDS, and the increased symptoms are expected to reduce HRQOL by 10% (such as the predicted HRQOL change by a statistical model derived from large-scale surveys, controlling for other validated covariates), then an observed increase of 15% suggests response shift. Thus, if change in cognitive appraisal were able to explain these systematic discrepancies, that was indicative of response shift. We decided to focus on the mental component score of the SF36v.2 rather than the physical component for this demonstration, because our prior preliminary analysis showed that it is more sensitive to response shift [19].

In order to provide a highly conservative test of response shift, we used an ordinary least square regression to control baseline mental composite score for a wide range of possible predictors, including demographics and personal history (e.g., history of hard drug use and involvement in the criminal justice system), baseline health status, baseline frame of reference, baseline sampling, standards, and combinatory algorithm, change in health status variables, changes in number of self-reported symptoms, and intervening events in care. Details on how these covariates are assessed can be found in [2,16]. Standardized residual scores controlling these predictors were computed and entered into the rpart analysis [20,21,22] to determine whether and how changes in cognitive appraisal could be used to explain these discrepancies.

2.7 Rpart model specifications

The rpart [20,21,22] model fitted the standardized residual MCS scores in SF36v2 [3] with 38 predictors representing the changes in appraisal variables between baseline and 6-months assessment: changes in 16 primary and 7 subsidiary goal content dimensions, 5 predictors on the sampling of experiences, 3 predictors on standards of comparisons, and 7 predictors on the combinatory algorithms. For ease of interpreting the magnitude of response shift, we divided our sample by the MCS discrepancy scores to three categories: 40% with the largest positive residuals (deemed "Positive" response shift), 40% with the largest negative residuals ("Negative" response shift), and 20% with residuals close to zero ("No Change").

We followed the general approach in rpart analysis: first grow a complex tree and then prune the tree back by cross-validation [20,21,23,24,25,26,27,28,29]. Feldesman [29] is a highly-accessible tutorial on the different statistical computations for continuous and categorical outcome variables; it also outlines a few default model specifications in the complex tree: 1) stopping rule for a terminal node (< 20 observations); 2) criterion for tree pruning ("cost-complexity parameter", CP=0.01); 3) validation by 10-fold cross-validation (1-SE rule for pruning by CP); 4) specification of priors (proportional to data counts) and 5) missing data are handled by surrogate splits. Details can be found in textbooks and are omitted here [15,28, 30,31].

2.7 Rpart model fit evaluation

Model performance was evaluated by a 3-class classification performance metric based on overall error rate in a confusion matrix, and also by pairwise area under the ROC curve analysis (AUC) [32,33]. Our three-class classification was separated into six binary comparisons

after the rpart classifier has been carried out. We calculated the AUC on "Positive" vs. "No Change" response shift, "Positive" vs. "Negative" response shift, and so on for all 6 pairwise ROCs. A single average AUC index was calculated, called the M function [33], to represent the overall model performance. We also entered the same 38 predictors in a multinomial logit model for comparison.

3. Results

3.1 Respondent characteristics

At this time of analysis, 619 individuals were recruited to this study, of which 443 were due for the 6 months assessments and 394 completed them (89%, follow-up data collection ongoing). Table 1 summarizes participant characteristics. Men and women were approximately evenly distributed, with diverse race and ethnicity backgrounds, low socioeconomic status, and an average age of 47.1 years and 11.6 years since the identification of HIV.

3.2 Response shift analysis using RPART

Figure 1 (a) shows the fullest rpart dendrogram, derived by accepting the default settings. The 10-fold cross-validation suggested pruning the tree back to only 9 terminal nodes (Figure 1 (b)). This was based on the 1 SE rule [20, 21], plotted in Figure 2, to find the least complex tree within 1 standard deviation of the minimal cross validation error. The pruned tree showed the lowest cross-validation error, beyond which tree complexity entailed no additional improvement.

Table 3 shows the confusion matrix of the 9-node tree and the model performance AUC measures of three alternative models. The 9-node tree made 243 correct classifications (62% accuracy, 95% CI = 39% – 80% by bootstrapping), which was superior to the 36% chance accuracy by the marginal 40-20-40 split. The pairwise AUC indexes show comparable performance between the 9-node rpart tree and the multinomial logit, with an overall AUC of 0.72. The 24-node tree consistently outperforms the pruned tree as well as the multinomial logit model. However, the cross-validation argued against it because of the low generalizability.

We now discuss the nine terminal nodes in Figure 1 (b) from left to right by interpreting the distinctions among groups which emerged in this analysis. The first group of 78 individuals in node 1 stood out because they reduced the salience of negative experiences by a moderately large amount. This group tended to have a high prevalence of positive discrepancies (47 out of 78). For the remaining 316 individuals in nodes 2 through 9, the salience of negative experiences in evaluating HRQOL was either maintained or increased ($>= -1.04$ sd). For succinctness we interpret non-large reduction in splits as roughly maintenance or possible increase. Reduction in salience of negative experiences alone was not sufficient to affect discrepancies in HRQOL. A combination of other cognitive variables comes into play. The second, third and fourth nodes were distinguished from the rest of the sample based on a moderate reduction in the extent to which they compared themselves to others. Group 4 represented a small subgroup that differed from groups 2 and 3 by a moderately large increase in goals related to solving problems associated with living situations and work. The majority of individuals in this small group demonstrated little or no discrepancy from expected change in psychological well-being, although three displayed negative discrepancies. For individuals in nodes 2 and 3, discrepancy was associated with moderate changes in goals related to independent functioning. Group 2 maintained or increased goals related to independence, which was associated with predominately positive response shift. Conversely, group 3 markedly reduced goals associated with independence, contributing to more negative psychological well-being than expected. It is noteworthy that groups 2, 3, and 4 were all

affected by changes in specific goals related to problems with work or with maintaining independence. Such changes in frame of reference interact with changes in standards of comparison to produce a range of response shifts.

Nodes 5 – 9 either maintained or increased their tendency to compare themselves with others, as well as the salience of negative feelings and experiences. On the far right node 9, the largest group of 145 individuals identified in this analysis sample, stood out from the others because of maintained or increased concerns about monetary obligations and other external demands. This combination, greater salience of negative feelings and comparison of self to others along with increased demands, was clearly associated with marked negative discrepancies in reported quality of life. However, for individuals without concerns about reducing demands and obligations, other factors came into play. Node 8 represented a group that had moderately high reduction in goals on maintaining relationships by accepting others and improving their own outlook. This group tended to report changes in psychological well-being close to values predicted by baseline factors, health changes and events in care.

Individuals in the remaining groups 5, 6, and 7 all tended to increase or sustain their concerns about maintaining relationships and achieving a positive outlook. Again moving in from the right, node 7 contained a preponderance of individuals with negative response shift. Interestingly, this group reported a marked decrease in goals related to preventing or avoiding interpersonal and monetary concerns. Conceivably, these individuals wanted to stave off certain problems at baseline but later realized that this was untenable by 6 months, thus contributing to a negative response shift.

Nodes 5 and 6 share many features, including the salience of negative experiences, a tendency to compare oneself to others, goals to maintain relationships, improve one's outlook, and avoid interpersonal and monetary problems, but not to reduce obligations or demands. The majority of these individuals, in node 5, tended to report more positive psychological well-being than expected as a result of efforts to manage communications during the interview. This factor, from the sampling experience domain reflects individuals' efforts to edit their responses by not complaining too much, by giving their first reaction, and by trying to convey the seriousness of their situation. Increasing or sustaining this response set was associated with positive discrepancies in well-being. Alternatively, individuals in node 6 had reduced or abandoned efforts to manage communication during the interview. The majority of individuals in node 6 demonstrated negative response shift.

4. Discussion

The rpart-derived model was useful in identifying aspects of response shift that was hard to detect through linear analysis. Rpart performed equally well as a multinomial logistic regression of the same predictors. Rpart provided a straightforward method that yielded more clinically interpretable results for identifying subgroups of response shift, which appeared to be mostly influenced by changes in emotion (e.g., 'negative experiences and feelings') and subjective norms ('comparing to others'). Invoking comparisons with others played an important role in explaining discrepancies in SF36v2-MOS change which only became apparent when examined in conjunction with the salience of negative events. Thereafter, frames of reference and individual concerns came into play in response shift, including 5 of the 16 primary goal factors. Our findings also shed light on the potential for individuals to sample and report experiences that affect their HRQOL selectively, to manage communication during the interview. Permitting the interplay of cognitive change variables provides additional, complementary information about QOL response shift.

The present findings have bearings on application of QOL measures and on understanding and meeting patients' needs. Individuals continually encounter new challenges and new opportunities, and factor these into how they self-evaluate their well-being. Individuals living with a chronic, life-threatening illness encounter many such challenges, and the stakes are high: An interpersonal conflict may interfere with an important source of social support; monetary or housing problems may strain an individual's limited resources; challenges to independence may invoke personal fears of premature debilitation and mortality. These challenges may be even greater in the Medicaid HIV/AIDS population. Additionally, we show that some individuals adapt to challenges by altering their cognitions, by relinquishing goals or modifying expectations for what they are seeking (e.g., node 9 by 'reducing practical & monetary obligations/demands' and node 7 by 'avoiding interpersonal & monetary concerns'). The distinction between nodes 7 and 9 is subtle, with node 9 emphasizing more on practical concerns such as reducing the financial obligations of paying bills. Individuals may selectively disengage from situations that they can no longer manage. These processes necessarily play out over time. It is not surprising that individuals' cognitive criteria for the appraisal of psychological well-being and distress is quite fluid.

Cognitive assessment provides a way to take these wide variations in quality of life appraisal into account [1]. We can use these methods to control response shift effects in evaluations of programs or treatments. For example, we might observe improvement in an individual's emotional well-being if we take into account that they are presently engaged in solving housing problems and in boosting independence (e.g., note 2). Similarly, apparent reduction in emotional well-being might be reinterpreted in light of an individual's selective emphasis on reducing practical and monetary demands (e.g., node 9). More fundamentally, it may be important to interpret the impact of disease and treatment on measures of cognitive appraisal. As our analysis demonstrates, there is a complex interplay among measures of appraisal and quality of life. It is important to understand when and how increased contact with the health system is associated with a sense of greater dependence, and when it is associated with increasing expectations and standards for self-evaluation. Change in the content and process of cognitive appraisal is a worthy patient-reported outcome domain in its own right.

Our results support the Rapkin and Schwartz model [1], in that cognitive variables helped to account for substantial HRQOL response shift in ways that were interpretable and consistent. However, several methodologic challenges remain. The QOL appraisal battery generates considerable, detailed descriptive data about the appraisal process. It is quite challenging to operationalize the process of describing the intermediaries of response shift, as we have attempted. There are inherent problems in CART methods [34], which originate from the fact that one predictor may win a particular split by only a small margin. This makes such splits somewhat arbitrary; errors cascade into subsequent splits, highlighting the importance of tree pruning by cross-validation [34]. New methodological developments are available [35,36,37, 38], and may be helpful in future research on response shift. We hope that our study will prompt further theoretical and empirical work to improve on the description of the response shift phenomena.

Acknowledgments

Bibliography

1. Rapkin B, Schwartz CE. Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift. Health Qual Life Outcomes 2004;2:14. [PubMed: 15023229]

2. Rapkin B, Weiss E, Chhabra R, Ryniker L, Patel S, Carness J, Adsuar R, Kahalas W, Delemarter C, Feldman I, Delorenzo J, Tanner E. Beyond satisfaction: using the Dynamics of Care assessment to better understand patients' experiences in care. Health Qual Life Outcomes 2008;6:20. [PubMed: 18331632]

3. Ware, JE.; Kosinski, MA.; Dewey, JE. How to Score Version 2 of the SF-36® Health Survey (Standard & Acute Forms). Lincoln, RI: Quality Metric Incorporated; 2000.

4. Rapkin, B. Personal goals and response shifts: Understanding the impact of illness and events on the quality of life of people living with AIDS. In: Schwartz, CA.; Sprangers, MAG., editors. Adaptation to Changing Health: Response Shift in Quality-of-Life Research. Washington, DC: American Psychological Association; 2000. p. 53-71.

5. Schwartz CE, Sprangers MAG. Methodological approaches for assessing response shift in longitudinal quality of life research. Soc Sci Med 1999;48:1531–1548. [PubMed: 10400255]

6. Sprangers MAG, Schwartz CE. Integrating response shift into health-related quality-of-life research: A theoretical model. Soc Sci Med 1999;48:1507–1515. [PubMed: 10400253]

7. Ubel PA, Loewenstein G, Jepson C. Whose Quality of Life? A commentary on exploring discrepancies between health state evaluations of patients and the general public. Qual Life Res 2003;12:599–607. [PubMed: 14516169]

8. Ubel PA, Loewenstein G, Schwarz N, Smith D. Misimagining the unimaginable: the disability paradox and health care decision making. Health Psychol 2005;24:S57–S62. [PubMed: 16045420]

9. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. J Chronic Dis 1978;31:697–704. [PubMed: 730825]

10. Schwartz CE, Andresen EM, Nosek MA, Krahn GL. Response shift theory: important implications for measuring quality of life in people with disability. Arch Phys Med Rehabil 2007;88:529–536. [PubMed: 17398257]

11. Weiden P, Rapkin B, Mott T, Zygmunt A, Goldman D, Horvitz-Lennon M, Frances A. Rating of medication influences (ROMI) scale in schizophrenia. Schizophr Bull 1994;20:297–310. [PubMed: 7916162]

12. O'Boyle CA, McGee H, Hickey A, O'Malley K, Joyce CR. Individual quality of life in patients undergoing hip replacement. Lancet 1992;339:1088–1091. [PubMed: 1349111]

13. McGee H, O'Boyle CA, Hickey A, O'Malley K, Joyce CRB. Assessing the quality of life of the individual: the SEIQoL with a healthy and a gastroenterology unit population. Psychol Med 1991;21:749–759. [PubMed: 1946863]

14. Sprangers MAG, van Dam F, Broersen J, Lodder L, Wever L, Visser MRM, Oosterveld P, Smets EMA. Revealing response shift in longitudinal research on fatigue: The use of the then-test approach. Acta Oncol 1999;38:709–718. [PubMed: 10522761]

15. Breiman, L.; Friedman, JH.; Olshen, RA.; Stone, CJ. Classification and Regression Trees. New York: Chapman Hall/CRC; 1984.

16. Patel S, Weiss E, Chhabra R, Ryniker L, Adsuar R, Carness J, Kahalas W, Delamarter C, Feldman IS, Delorenzo JP, Tanner E, Rapkin B. The Events in Care Screening Questionnaire (ECSQ): a new tool to identify needs and concerns of people with HIV/AIDS. AIDS Patient Care STDS 2008;22:381–393. [PubMed: 18366327]

17. Cohen, J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

18. Cohen J. A power primer. Psychol Bull 1992;112:155–159. [PubMed: 19565683]

19. Li, Y.; Rapkin, B. HIV/AIDS Patients' Quality of Life Appraisal Depends on Their Personal Meaning Of Quality of Life and Frame of Reference; Qual Life Res. 2006 [assessed January 15, 2009]. p. A-36Available online at: http://www.springerlink.com/content/h48815020815g7w6/fulltext.pdf

20. Therneau, TM.; Atkinson, EJ. An Introduction to Recursive Partitioning Using The Rpart Routines. Rochester, MN: Mayo Foundation; 1997.

21. Venables, WN.; Ripley, BD. Modern Applied Statistics with S. New York: Springer Science +Business Media; 2002.

22. R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: The R Foundation for Statistical Computing; 2008.

23. Martin MA, Meyricke R, O'Neill T, Roberts S. Mastectomy or breast conserving surgery? Factors affecting type of surgical treatment for breast cancer--a classification tree approach. BMC Cancer 2006;6:98. [PubMed: 16623956]

24. Gruenewald TL, Mroczek DK, Ryff CD, Singer BH. Diverse pathways to positive and negative affect in adulthood and later life: an integrative approach using recursive partitioning. Dev Psychol 2008;44:330–343. [PubMed: 18331126]

25. Radespiel-Troger M, Rabenstein T, Schneider HT, Lausen B. Comparison of tree-based methods for prognostic stratification of survival data. Artif Intell Med 2003;28:323–341. [PubMed: 12927339]

26. Sedrakyan A, Zhang H, Treasure T, Krumholz HM. Recursive partitioning-based preoperative risk stratification for atrial fibrillation after coronary artery bypass surgery. Am Heart J 2006;151:720–724. [PubMed: 16504639]

27. Nicolucci A, Carinci F, Ciampi A. Stratifying patients at risk of diabetic complications: an integrated look at clinical, socioeconomic, and care-related factors. SID-AMD Italian Study Group for the Implementation of the St. Vincent Declaration. Diabetes Care 1998;21:1439–1444. [PubMed: 9727888]

28. Hastie, T.; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer; 2001.

29. Feldesman MR. Classification trees as an alternative to linear discriminant analysis. Am J Phys Anthropol 2002;119:257–275. [PubMed: 12365038]

30. Zhang, H.; Singer, B. Recursive Partitioning in The Health Sciences. New York: Springer-Verlag; 1999.

31. Clark, LA.; Pregibon, D. Tree-Based Models. In: Chambers, JM.; Hastie, TJ., editors. Statistical models in S. New York: Chapman and Hall/CRC; 1991.

32. Patel, AC.; Markey, MK. Comparison of three-class classification performance metrices: a case study in breast cancer CAD. In: Eckstein, MP.; Jiang, Y., editors. Proceedings of the SPIE, Medical Imaging 2005: Image Perception, Observer Performance, and Technology Assessment; Bellingham, WA: SPIE; 2005. p. 581-589.

33. Hand DJ, Till RJ. A simple generalisation of the area under the ROC curve for multiple class classification problems. Machine Learning 2001;45:171–186.

34. Radespiel-Troger M, Hothorn T, Pfahlberg AB, Gefeller O. Re: "Applying recursive partitioning to a prospective study of factors associated with adherence to mammography screening guidelines". Am J Epidemiol 2006;164:400–401. [PubMed: 16809428]

35. Breiman L, Cutler A. Random Forest. Machine Learning 2001;45:5–32.

36. Breiman L. Bagging Predictors. Machine Learning 1996;24:123–140.

37. Hothorn T, Lausen B, Benner A, Radespiel-Troger M. Bagging survival trees. Stat Med 2004;23:77–91. [PubMed: 14695641]

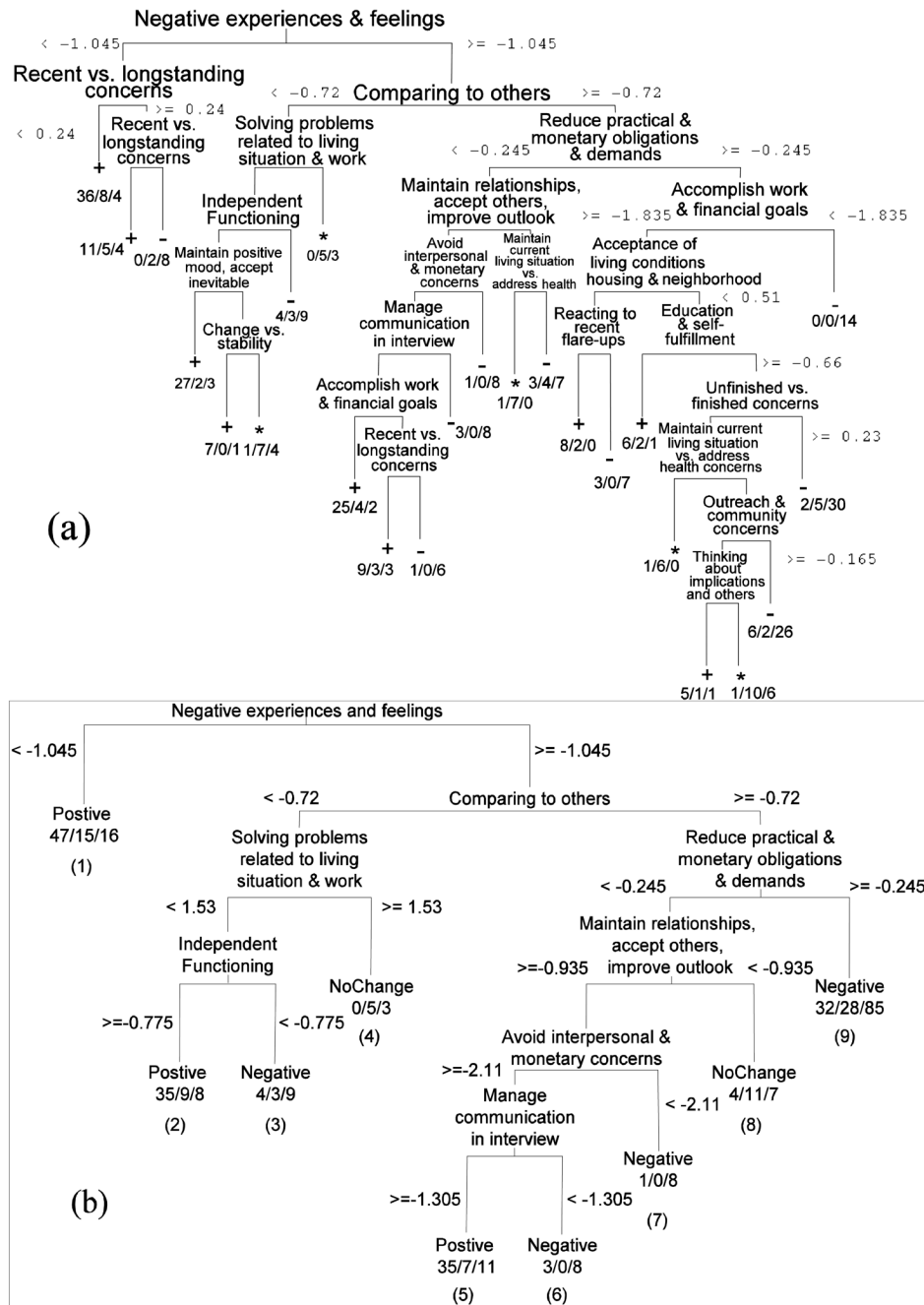38. Hothorn, T.; Hornik, K.; Zeileis, A. party: A Laboratory for Recursive Part(y)itioning. 2006. (www.r-project.org).

Figure 1.
Full rpart tree with 24 terminal nodes (a) and the pruned tree with 9 terminal nodes (b). To minimized clutter in (a), the most prevalent outcomes in the terminal nodes are represented in symbols, for positive discrepancy (marked with the + sign), negative discrepancy (−) and no discrepancy (*). Many splitting criteria in (a) are omitted.
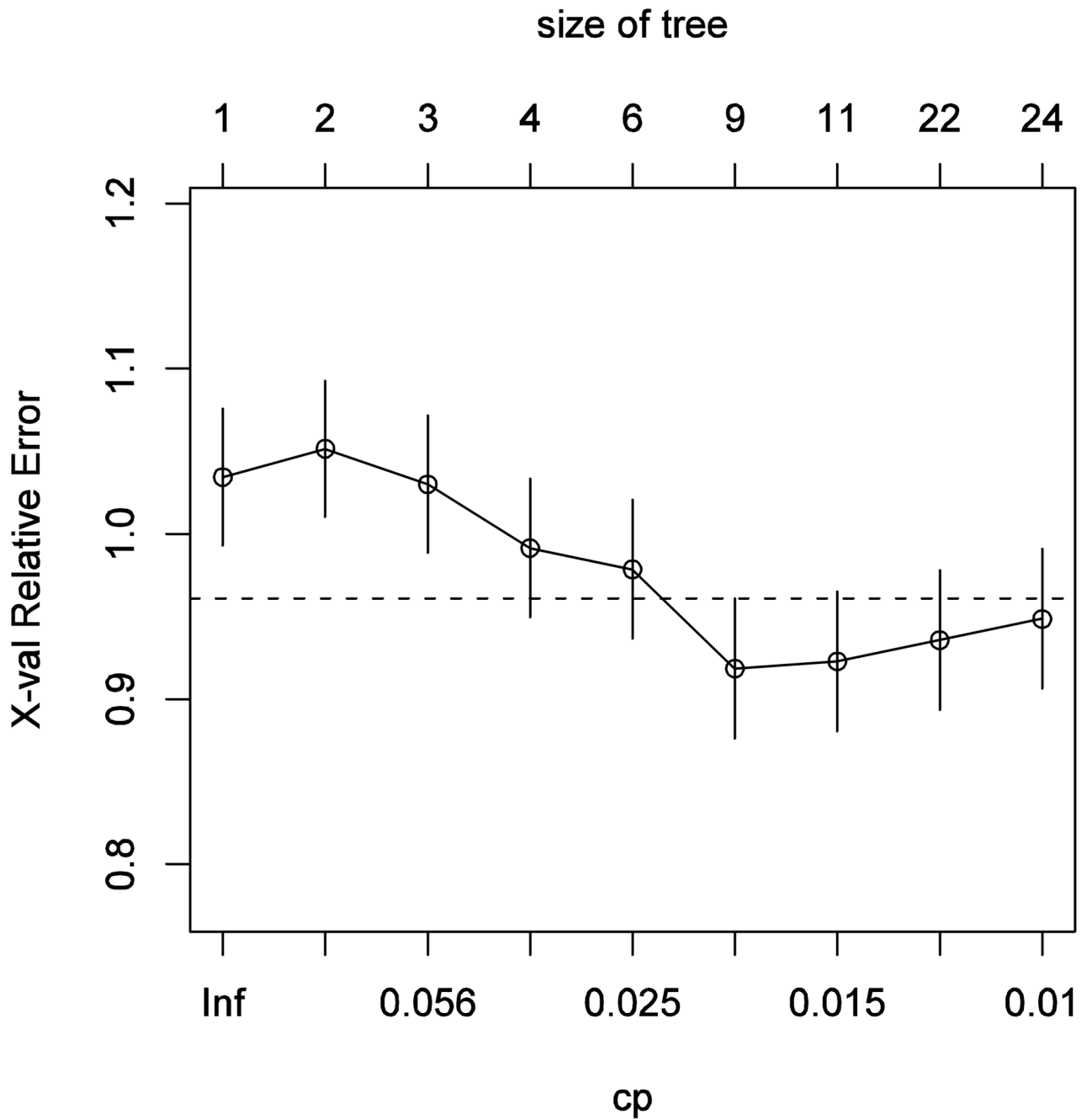
Figure 2.
Graphical output of the plotcp() command in rpart. The horizontal dotted line represents the 1 SE rule---the cutoff cross-validation error statistic at 1 standard deviation above the minimal cross-validation error. The tree with 9 terminal nodes is considered the desired size for pruning because it entails the lowest cross-validation before additional complexity in the tree are accompanied by higher cross-validation errors.

Table 1

Participant Characteristics.

Characteristic	n = 619
Sex (% male)	328 (53%)
Age in years	47.1 years (sd = 8.5)
Time since HIV identification (years)	11.6 years (sd = 5.7)
Marital status/domestic partner	409 (66%)
Sexual orientation (% heterosexual)	452 (73%)
Race	
African descent	359 (58%)
Anglo	31 (5%)
Latino	186 (35%)

Note: Numbers are persons and percentages unless otherwise noted.

Table 2

Primary and subsidiary goal content dimensions.

	Primary goal content dimensions	Example verbatim response
1	Maintain relationships, accept others, improve outlook	'I want to keep my marriage the same as it is now' (865.6)
2	Solve problems with HIV treatment	'I want to solve my HIV problems' (887.2)
3	Prevent money and housing problems, reduce worries and obligations	'I want accomplish getting better housing' (578.3)
4	Solve problems related to living situation and work	'I want to solve problems in my living situation, I need to move' (53.4)
5	Address physical and emotional health problems	'I want to prevent or avoid pain, both physical and emotional' (313.4)
6	Learn to live with HIV diagnosis and maintain current treatment arrangements	'I want to accept that I've got to live with HIV and accept as they are' (325.6)
7	Avoid interpersonal and monetary concerns	'I want to avoid getting into fights and arguments' (1849.8)
8	Maintain a positive mood, learn to accept the inevitable	'I want to keep the same my attitude in thinking positively' (2116.9)
9	Accomplish work and financial goals	'I want to be more financially stable' (177.1)
10	Avoid work-related problems	'I don't want anybody at work to know my HIV status' (137.8)
11	Reduce practical and monetary obligations and demands	'I want to reduce and responsibilities of credit card bills' (1808.9)
12	Maintain current living situation (versus address health problems)	'I want to keep the same my living arrangements with my wife and family' (2331.10)
13	Acceptance, resignation to health and mood problems	'I want to accept that nobody lives forever' (783.4)
14	Maintain and accept current work and monetary situation	'I want to accept that I have to work even with my health problems' (2127.7)
15	Acceptance of living conditions, housing and neighborhood	'I want to keep living in my neighborhood' (2565.6)
16	Concerns about HIV Prevention	'I want to prevent/avoid infecting other people with HIV' (209.1)

Subsidiary goal content dimensions

1	Outreach and community concerns	'I want to accomplish more in outreach work' (210.1)
2	Social, religious and discrimination concerns	'I want to turn to God more' (1076.7)
3	Travel and leisure versus chores	'I want to travel to my country' (1707.7)
4	Independent Functioning	'I want to keep the same my independence' (2053.9)
5	Education and self-fulfillment	'I want to accomplish the goal of finishing school' (367.8)
6	Legal and immigration concerns	'I want to solve the problem of my immigration status' (2261.10)
7	Substance use	'My main problem to solve is to break the methadone habit' (781.1)

Selected verbatim responses are identified by two numbers so that (781.1) represent statement number 781 assigned to rater 1.

Table 3

Factor loadings of QOL appraisal scales in Sampling Experiences.

"When you responded today, how much did you…"	1 [b]	2	3	4	5
	14% [a]	14%	13%	11%	9%
Find yourself thinking about the worst moments?	0.78				
Focus on HIV/AIDS?	0.67				
Consider things that you'd only think about for an interview like this?	0.57				
Try to give your first reaction to the questions?		0.77			
Try not to complain too much?		0.74			
Try to communicate the seriousness of your situation?		0.64			
Try to remember everything relevant over the past 3 months?			0.78		
Think about how things have been going over past few days?			0.70		
Emphasize the positive as much as possible?	−0.42		0.51		
Consider your relationships with family/friends?				0.83	
Take into account what your doctor has told you about your health?				0.59	0.46
Think about the future?					0.82
Balance the positives with the negatives?		0.42		0.44	
Recall recent episodes or flare-ups?	0.44		0.43		

Note: loadings < 0.40 were omitted.

[a] Variance accounted for by the eigen values associated with the latent factors.

[b] Factor 1: "Focusing on worst moments re illness"; Factor 2: "Formulating responses to manage interview"; Factor 3: "Recalling recent events"; Factor 4: "Considering interactions with family and others"; and Factor 5: "Contemplating the future".

Table 4

Factor analysis and factor loadings on focus of comparisons.

"When rating your health and well-being today, how much did you compare yourself to …"	1	2	3
	23%	23%	20%
Others you know who are living now with HIV/AIDS?	0.89		
People whose health does not limit them in any way?	0.89		
Your ideal: your dream of perfect health?		0.82	
The kind of life that you are really working for?		0.82	
A time in your past before you had HIV?			0.88
Most people your age?		0.40	0.63
The way that the people in your life see you?			0.45
The things your doctor told you would happen?			0.41

1. Comparing oneself to others with HIV and with no health limitations

2. Comparing oneself to personal ideals or desired goals

3. Comparing oneself to one's past and to age-related norms.

Table 5

Factor analysis and factor loadings on salience of experiences.

"When you answered today, did you think more about …"	1	2	3	4	5	6	7
	14%	**12%**	**10%**	**9%**	**8%**	**8%**	**7%**
Things that are disappointing to you?	0.78						
How hard it has been?	0.65						
Things that make you feel worried?	0.64						
Things that you are proud of?		0.74					
How well you are doing?		0.70					
Things that make you feel calm?		0.70					
Big changes?			0.78				
The way things usually are?			−0.76				
Things settled to your satisfaction?				0.83			
Things that are unfinished?				−0.74			
What is important to others?					0.76		
What is important to you?					−0.77		
Long time concerns?						0.79	
Recent concerns?	0.40					−0.72	
What you do on your own?							0.90
The help you need from other people?	0.46						−0.52

1. Negative experiences, feelings & worries

2. Sources of satisfaction

3. Change vs. routine

4. Settled vs. unfinished concerns

5. Things important to others vs. self

6. Long-time vs. recent concerns

7. Independence vs. help from others

Table 6

Classification and misclassification results of the pruned rpart tree with 9 terminal nodes as well as model performance indexes for three alternative models.

		Predicted Response Shift			Total
		Positive	No Change	Negative	
Residual	Positive	117	4	40	161
Changes in	No Change	31	16	31	78
HRQOL	Negative	35	10	110	110
	Total	183	30	181	394

	Pairwise AUC						Overall AUC
	1	2	3	4	5	6	
Rpart 24-node tree	0.83	0.9	0.87	0.83	0.90	0.87	0.86
Rpart 9-node tree	0.70	0.77	0.69	0.70	0.76	0.69	0.72
Multinomial logit	0.71	0.75	0.71	0.71	0.75	0.71	0.72

1: No Change *vs.* Positive

2: Negative *vs.* Positive

3: Negative *vs.* No Change

4: Positive *vs.* No Change

5: Positive *vs.* Negative

6: No Change *vs.* Negative