Measures of Relative Importance for Health-Related Quality of Life

Tolulope T. Sajobi[1], Lisa M. Lix[1,2], Ian Clara[2,3], John Walker[2,4,5], Lesley A. Graff[2,4], Patricia Rawsthorne[2], Norine Miller[2], Linda Rogala[2], Rachel Carr[2,5], Charles N. Bernstein[2, 6]

[1]School of Public Health, University of Saskatchewan, Saskatoon, CANADA

[2]University of Manitoba Inflammatory Bowel Disease Clinical and Research Centre, University of Manitoba, Winnipeg, CANADA

[3]Department of Psychology, University of Manitoba, Winnipeg, CANADA

[4]Department of Clinical Health Psychology, University of Manitoba, Winnipeg, CANADA

[5]Department of Community Health Sciences, University of Manitoba, Winnipeg, CANADA

[6]Department of Internal Medicine, University of Manitoba, Winnipeg, CANADA

Corresponding Author:
Lisa Lix
School of Public Health
University of Saskatchewan
107 Wiggins Road
Saskatoon, SK CANADA S7N 5E5
Phone: +1 306-966-1617
Fax: +1 306-966-7920
e-mail: lisa.lix@usask.ca

Manuscript Word Count: 3851

Abstract Word Count: 197

Abstract

Purpose: In health-related quality of life (HRQOL) studies, data are often collected on multiple domains for two or more groups of study participants. Quantitative measures of relative importance, which are used to rank order the domains based on their ability to discriminate between groups, are an alternative to multiple tests of significance on the group differences. This study describes relative importance measures based on logistic regression (LR) and multivariate analysis of variance (MANOVA) models. **Methods:** Relative importance measures are illustrated using data from the Manitoba Inflammatory Bowel Disease (IBD) Cohort Study. Study participants with self-reported active (n = 244) and inactive (n = 105) disease were compared on 12 HRQOL domains from the Inflammatory Bowel Disease Questionnaire (IBDQ) and Medical Outcomes Study 36-item Short-Form (SF-36) Questionnaire. **Results:** All but two relative importance measures ranked the IBDQ bowel symptoms and emotional health domains as most important. **Conclusions:** MANOVA-based importance measures are recommended for multivariate normal data and when group covariances are equal, while LR measures are recommended for non-normal data and when the correlations among the domains are small. Relative importance measures can be used in exploratory studies to identify a small set of domains for further research.

Key Words: discriminant analysis; health-related quality of life; inflammatory bowel disease; logistic model; multivariate analysis; relative importance

Abbreviations

ADRC = adjusted discriminant ratio coefficient
API = adjusted Pratt's Index
BP = bodily pain
BS = bowel symptoms
DDA = descriptive discriminant analysis
DRC = discriminant ratio coefficient
EH = emotional health
GH = general health
HRQOL = health-related quality of life
IBD = inflammatory bowel disease
IBDQ = Inflammatory Bowel Disease Questionnaire
LR = logistic regression
MANOVA = multivariate analysis of variance
MH = mental health
OLS = ordinary least squares
PF = physical functioning
PI = Pratt's index
RE = role emotional
RP = role physical
RW = relative weight
RRW = rescaled relative weight
SDFC = standardized discriminant function coefficient
SF = social functioning
SF-36 = 36-item Short Form Questionnaire
SLRC = standardized logistic regression coefficient
SS = systemic symptoms
VT = vitality

Measures of Relative Importance for Health-Related Quality of Life

1. Introduction

In studies about health-related quality of life (HRQOL), data are often collected on multiple domains, such as physical function, social health, and emotional health, for two or more groups of study participants (e.g., treatment and control groups) [1]. A single overall test of group differences can be obtained using a multivariate procedure, such as multivariate analysis of variance (MANOVA) [2]. However, researchers are often interested in identifying the domain(s) on which group differences exist [3, 4]. While multiple tests of significance for the group differences could be performed, using an appropriate multiple testing procedure to control the overall probability of a Type I error, an alternative approach is to adopt a measure of relative importance. Relative importance measures can be used to rank order the domains based on their ability to discriminate between groups. They enable researchers to make statements like "X was the most important HRQOL domain amongst those studied". The measures have a number of uses, such as developing parsimonious statistical models, identifying the domains to target in clinical interventions, or identifying the domain(s) on which a treatment or intervention has the greatest effect.

Although relative importance measures have been used in other disciplines [5-6], they may not be familiar to researchers who investigate HRQOL. The measures are rarely discussed in textbooks and research on this topic has primarily appeared in statistical journals. As well, given that no single measure is uniformly recommended and the measures will not always produce consistent results [7], implementing a relative importance analysis may not be straightforward.

The purpose of this study is to describe measures to quantify the relative importance of HRQOL domains and examine their properties under the following data characteristics that are

likely to be encountered in HRQOL studies [1, 8-10]: (a) non-normal data, (b) between-group variance heterogeneity, (c) collinearity of domains, and (d) missing data. The measures are illustrated using data from a cohort study about inflammatory bowel disease (IBD).

2. Description and Comparison of Relative Importance Measures

Relative importance measures for studies involving two groups include: (a) standardized logistic regression coefficients (SLRCs) [11], (b) Pratt's index for logistic regression [12], (c), dominance analysis [13], (d) relative weights (RWs) [14], (e) standardized discriminant function coefficients (SDFCs) [15], (f) discriminant ratio coefficients (DRCs) [16], and (g) *F*-to-remove statistics [16]. The first four measures, which are obtained from the LR model, treat the domains as explanatory variables, while the last three, which are based on the MANOVA model, treat them as outcome variables. Effect size measures or *p*-values are sometimes used to assess relative importance [17]. However, effect size measures like Cohen's *d* [18], describe absolute importance and do not account for the correlations amongst the variables [17-19]. A *p*-value provides a measure of statistical importance, that is, the probability of a result under the assumption that the null hypothesis is true.

We focus on measures of relative importance for studies involving two independent groups of study participants. Measures for three or more groups, which have been developed using the MANOVA model, are discussed at the end of this section.

To begin, assume that data are available on $m \geq 2$ HRQOL domains for N study participants, with n_1 study participants from group 1 and n_2 study participants from group 2 ($n_1 + n_2 = N$).

2.1 Measures Based on the LR Model

The LR model is [20]

$$\log\left(\frac{p_i}{1-p_i}\right) = \mathbf{X}_i\boldsymbol{\beta}, \quad (1)$$

where $p_i = \Pr(y_i = 1 | \mathbf{X}_i)$ is the probability the ith study participant ($i = 1, \ldots, N$) is a member of group 1 conditional on the explanatory variables (e.g., HRQOL domains), $\mathbf{X}_i$ is a vector of dimension ($m + 1$) where the first element is equal to one, and $\boldsymbol{\beta}$ is a ($m + 1$) vector of regression coefficients to be estimated, with the first element equal to the model intercept, β_0. Details about the estimated coefficients are provided in Appendix A.

The use of SLRCs to rank order the domains has been proposed in several papers [21-23]. Most methods to calculate standardized coefficients are partial methods that do not account for variation in the grouping variable [21, 24]. Fully standardized coefficients, which are recommended for assessing relative importance, were proposed by Mernard [11, 21]. The kth SLRC is

$$\hat{\beta}_k^* = \hat{\beta}_k s_{x_k} R / s_{\text{logit}(\hat{p})}, \quad (2)$$

where $\hat{\beta}_k$ is the estimated coefficient, s_{x_k} is the sample standard deviation for the kth domain, R is the square root of the coefficient of determination (i.e., R^2), and $s_{\text{logit}(\hat{p})}$ is the sample standard deviation of the logit of the predicted probabilities (i.e., $\hat{p}_i$s). While several formulae for R^2 have been proposed for the LR model, the most common formulae are based on the log of the likelihood function and ordinary least squares (OLS) regression of the dependent variable values on the $\hat{p}_i$s [13, 24-26]. The SLRCs usually range in value from -1 to +1, although values outside this range are possible. Larger absolute values indicate greater relative importance.

Pratt's index was first proposed for the OLS regression model [27-28] and then extended to the LR model [12]. Relative importance of the kth domain is the proportion of R^2 explained by it,

$$d_k = \frac{\hat{\beta}_k^* \hat{\rho}_k}{R^2}, \tag{3}$$

where $\hat{\rho}_k$ is the estimated correlation between the kth domain and the logit of the $\hat{p}_i$s, and $\hat{\beta}_k^*$ and R are defined in equation 2. The index generally ranges in value from zero to one with values greater than $1/2m$ indicating meaningful importance [29]. Small negative values between $-1/2m$ and zero can be set to zero, while large negative values indicate collinearity or suppression. The latter arises when a domain makes little or no direct contribution to the prediction of the outcome variable, but contributes indirectly through another domain. Suppressor variables have large negative values of Pratt's index but their SLRC values are usually similar in magnitude to the coefficients of non-suppressor variables. Potential suppressor or collinear variables should be excluded from the analysis and an adjusted Pratt's index (API) computed,

$$d_s^* = \frac{d_s}{\sum_{k \in S^c} d_k}, \qquad s \in S^c \tag{4}$$

where S^c is the set of indices corresponding to the non-suppressor domains. The variance inflation factor, an index that measures the increase in the variance of a regression coefficient due to collinearity [30], and the correlation matrix are useful for identifying collinear domains. Further details about Pratt's index are provided in Appendix A.

Dominance analysis was developed for assessing relative importance in the OLS model [31] and later extended to the LR model [13]. A dominance weight is the average change in R^2 when one domain is added to all possible subsets of the other domains. For example, given a model with three domains denoted by X_1, X_2, and X_3, the dominance weight for X_1 is obtained by computing the average change in R^2 when X_1 is added to regression models containing (a) the intercept only, (b) X_2 only, (c) X_3 only, and (d) X_2 and X_3. A larger weight indicates greater importance. Previous research has shown that the choice of R^2 statistics does not result in

appreciable differences in the dominance weights.

RW analysis was first proposed for the OLS model [32] and later extended to the LR model [14]. Like the previous measures of relative importance, the RWs are based on the model R^2, but a more complex set of computations is involved. Let $\mathbf{X}$ be a $N \times m$ matrix of standardized scores (i.e., mean = 0; variance = 1) for the domains. Then, $\mathbf{X}$ can be decomposed as,

$$\mathbf{X} = \mathbf{P}\Delta\mathbf{Q}^{\mathrm{T}}, \tag{5}$$

where $\mathbf{P}$ is a $N \times m$ matrix consisting of m eigenvectors of $\mathbf{X}\mathbf{X}^{\mathrm{T}}$, $\mathbf{Q}$ is a $m \times m$ matrix of the eigenvectors of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$, Δ is a $m \times m$ diagonal matrix based on the square roots of the eigenvalues of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$, and $^{\mathrm{T}}$ is the transpose operator. Let $\mathbf{Z} = \mathbf{P}\mathbf{Q}^{\mathrm{T}}$. The vector of RWs is,

$$\mathbf{w} = \Lambda^2\hat{\boldsymbol{\beta}}^{*2}, \tag{6}$$

where $\Lambda = (\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{X}$, the $m \times m$ matrix obtained by regressing the original domains on the orthogonal domains, $\hat{\boldsymbol{\beta}}^{*2}$ is a $m \times 1$ vector in which each element of $\hat{\boldsymbol{\beta}}^{*}$ is squared, and $\hat{\boldsymbol{\beta}}^{*}$ is the $m \times 1$ vector of SLRCs (see equation 2). The RWs sum to R^2. A rescaled RW (RRW), which is easier to interpret, is the proportion of R^2 explained by a domain. Larger RWs and RRWs indicate greater importance.

2.2 Measures Based on the MANOVA Model

MANOVA-based measures of relative importance are obtained from descriptive discriminant analysis (DDA) and stepwise MANOVA procedures [13, 16, 33]. DDA identifies the linear combination of domains that maximally separates the groups [34]. Let $\mathbf{X}_{ij}$ represent the vector of domain scores for the ith study participant in the jth group ($i = 1,\ldots, n_j$; $j = 1, 2$), $\overline{\mathbf{X}}_j$ is the vector of domain means for the jth group, and $\overline{\mathbf{X}}$ is the vector of overall means. The vector of discriminant function coefficients, $\mathbf{a}$, is estimated by

$$\hat{\mathbf{a}} = \mathbf{S}^{-1}(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2), \quad (7)$$

where **S** is the pooled sample covariance matrix, and $\hat{\mathbf{a}}^{\mathrm{T}}\mathbf{S}\hat{\mathbf{a}} = 1$. The coefficient for the *k*th domain, which corresponds to the *k*th element of $\hat{\mathbf{a}}$, has been shown to be mathematically equivalent (but not always numerically equal) to the *k*th LR coefficient [35]. Details about discriminant function coefficients are provided in Appendix A.

The SDFC for the *k*th domain, denoted by a_k^*, is the product of the discriminant function coefficient and the standard deviation for the *k*th domain. SDFCs can be positive or negative, and the absolute magnitude determines relative importance.

DRCs have been recommended by some researchers instead of SDFCs [16, 33]. The *k*th DRC is given by

$$q_k = a_k^* f_k, \quad (8)$$

where f_k is the *k*th structure coefficient, the correlation between the *k*th domain and the discriminant function. DRCs generally range in value from zero to one, with larger values indicating greater importance. Similar to Pratt's index, a DRC can have a negative value, which may be indicative of collinearity or suppression. In MANOVA, suppression occurs when a domain makes little or no direct contribution to group separation on its own but contributes indirectly through another domain. Adjusted DRC (ADRC) statistics can be produced in a similar way to API statistics (see equation 4).

The *F*-to-remove statistic [36] is obtained by conducting *m* MANOVA tests, each time removing one domain from the analysis. For the *k*th domain, the statistic can be defined as

$$F_{(k)} = k_2\left((D^2 - D_{(k)}^2)/(k_3 + D_{(k)}^2)\right), \quad (9)$$

where $k_2 = n_1 + n_2 - m$, $k_3 = (n_1 + n_2)(n_1 + n_2)/(n_1 n_2)$, $D^2 = (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2)^{\mathrm{T}}\mathbf{S}^{-1}(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2)$, and

$D^2_{(k)}$ represents the value of D^2 when the *k*th domain is omitted. Appendix A provides additional details. *F*-to-remove statistics have a lower bound of zero, but no upper bound. Relative importance is assessed by the magnitude of the *F*-to-remove-statistic, with the most important domain yielding the largest statistic [10, 15].

2.3 Measures for Three or More Groups

For studies involving three or more independent groups, only MANOVA-based measures of relative importance (i.e., SDFCs, DRCs, and *F*-to-remove statistics) can be computed. They are calculated from *c* sets of linear discriminant functions coefficients, where $c = min\ (m, g - 1)$ and *g* is the number of groups [37]. Weighted SDFCs and DRCs are used to assess importance, where the weights are the eigenvalues for each set of discriminant functions coefficients [15]. The SDFC for the *k*th domain is

$$a_k^{**} = \sum_{l=1}^{c} \lambda_l a_{lk}^{*}, \tag{10}$$

where λ_l is the eigenvalue that corresponds to the *l*th eigenvector ($l = 1, \ldots, c$) of $\mathbf{E}^{-1}\mathbf{H}$ (see Appendix A for formulae for **E** and **H**), and a_{lk}^{*} is the SDFC for the *l*th eigenvalue and *k*th domain. Similarly, the DRC for the *k*th domain is

$$q_k^{*} = \sum_{l=1}^{c} \lambda_l a_{lk}^{*} f_{lk}, \tag{11}$$

where f_{lk} is the structure coefficient for the *k*th domain and *l*th eigenvector. The *F*-to-remove statistic is readily extended to multi-group designs using equation 9 [36].

2.4 Choosing a Relative Importance Measure

The choice of a relative importance measure will depend, in part, on the characteristics of the data. The LR model assumes a linear relationship between the logit of the p_is and each of the

domains. MANOVA-based measures of relative importance rest on the assumptions of a multivariate normal distribution and homogeneity (i.e., equality) of group covariance matrices. When the assumptions of the MANOVA model are satisfied, it has greater statistical power to discriminate between groups than the LR model [27]. However, it is not known if this difference in power affects the ranking of the domains in a relative importance analysis. Finch and Laking [38] showed, however, that under assumption violations, relative importance measures based on SDFCs resulted in an incorrect rank ordering of the variables. When these assumptions are violated, measures based on the LR model should be selected.

Collinearity and suppression present a challenge in assessing relative importance. SLRCs, Pratt's index, and DRCs are sensitive to these data characteristics and therefore should not be adopted when the correlations amongst the domains are moderate to large in size. Dominance analysis and RW analysis are the least sensitive to the correlations.

Both LR and MANOVA models result in casewise deletion of observations with missing values. Casewise deletion can result in biased estimates of relative importance when the mechanism of missingness is not random [39-40]. While an imputation method could be adopted when there are missing values, there is no optimal method to control bias in parameter estimates. The choice of imputation methods will depend on the characteristics of the data.

Computational ease is also a consideration when adopting a relative important measure. While LR and MANOVA models can be implemented using existing statistical software such as SAS [41], RW and dominance analyses require a number of additional computations. In particular, dominance analysis requires more computational resources than other methods because of the number of regression models that are fit to the data. For example, for five variables, 31 separate regression models are required to calculate the dominance weights, while

1023 regression models are required for 10 variables.

3. Numeric Example

Measures of relative importance are illustrated using data from the Manitoba IBD Cohort Study, a prospective longitudinal study, initiated in 2002, of patients who were recently diagnosed with Crohn's disease or ulcerative colitis [42-43]. Ethics approval for the Cohort Study was obtained from the University of Manitoba Health Research Ethics Board.

A total of 388 participants were initially enrolled in the Cohort Study. Data are collected using standardized self-report instruments or interviews conducted at six-month intervals. In this example, we focus on measures for distinguishing between study participants with active and inactive disease using data collected at the baseline measurement occasion. Disease activity was assessed using self-reported IBD symptom persistence in the previous six months based on a question with a six-point response format, which was subsequently dichotomized. This measure has been validated and shown to have good concordance with clinical measures of disease activity [44].

3.1 Study Measures

HRQOL data was collected using the Inflammatory Bowel Disease Questionnaire (IBDQ) [45] and the Medical Outcomes Study 36-item Short-Form (SF-36) Questionnaire [46]. The IBDQ adopts a Likert response scale and encompasses four domains: bowel symptoms, emotional health, social function, and systemic symptoms. The average score on each domain ranges from one to seven, with higher scores indicating better HRQOL [47]. The SF-36 encompasses eight domains, role physical, bodily pain, physical functioning, general health, role emotional, mental health, vitality, and social functioning. The domain scores are scaled to range in value from zero (poor health) to 100 (good health).

A total of 356 participants provided data at the baseline occasion. Participants with missing values on disease activity or HRQOL domains constituted 2.0% of the sample. The relative importance analysis was carried out for the 349 study participants with complete data on all study measures.

3.2 Statistical Analysis

Study participants were initially described on a variety of socio-demographic and disease characteristics. Domain scores were summarized using means and standard deviations. The Shapiro-Wilk [48] test of normality was computed for the 12 domains, along with descriptive measures of skewness and kurtosis.

All relative importance measures were computed with the exception of dominance analysis, which is impracticable to use because it would require fitting 4095 regression models to the data. The signed values of Pratt's index and the DRCs are reported; other importance measures were reported using absolute values. A rank score was assigned to each domain for each measure, with a value of one representing the most important domain. Ties in ranks were resolved by assigning mid-ranks [15]. All analyses were conducted using SAS software [41].

SAS syntax to implement all relative importance measures is provided in the supplementary documentation and is illustrated in that documentation using a small clinical dataset.

3.3 Results

Demographic and disease characteristics of the study participants are reported in Table 1. Descriptive statistics for the 12 domain scores are provided in Table 2. There is little evidence of between-group covariance heterogeneity on the domains. While the Shapiro-Wilk test of normality was statistically significant for all domains ($p < 0.001$), univariate skewness values ranged from -1.40 to 0.91 and kurtosis ranged from -0.90 to 1.20 (skewness and kurtosis for

normal distribution = 0), indicating small to modest departures from a normal distribution.

The correlations amongst the domains are reported in Appendix B. Using Cohen's [18] effect size criterion, approximately half of the correlations are large (r = .50 or higher) or moderate (.30 ≤ r ≤ .49) in size.

The relative importance measures are reported in Table 3, along with the conventional two-group (i.e., pooled) t-tests for group differences. If each of the t-tests were conducted at the α = .05/12 = .004 significance level to control the overall probability of a Type I error, all the domains would result in statistically significant differences between the two groups. The SLRCs range in absolute value from 0.015 to 0.463. For Pratt's index, eight of the 12 values are positive. Of the four negative values, those for systemic symptoms, mental health and role emotional domains are large (i.e., < -1/2m = -0.042) and have moderate to large SLRCs, indicating that they are potential suppressor variables. The index value for the role physical domain is small and is therefore set to zero. The API was computed after removing the three domains with large negative values. RRWs for the domains range from 0.002 to 0.369. The SDFCs range in absolute value from 0.027 to 0.587. While the DRCs range from -0.072 to 0.542, only seven domains have positive values. Of the five domains with negative DRC values, only mental health and role emotional have large negative values and large SDFC values, suggesting they are suppressor variables. ADRC statistics were calculated after removing these two domains and setting the small negative values to zero. Finally, the F-to-remove statistics exhibited substantial variation and ranged from 0.011 to 14.334.

Table 4 and Figure 1 contain the rank scores for each of the relative importance measures. All measures, except the F-to-remove statistics and RRWs, were consistent in ranking the disease-specific IBDQ bowel symptoms and emotional health domains as most important. The

F-to-remove statistics identified the IBDQ systemic symptoms and SF-36 general health domains as the most important domains, while the RRWs resulted in the role physical and social functioning domains receiving the highest ranks. The SF-36 physical functioning domain was ranked as the third most important domain by the *F*-to-remove and RRW statistics, while this domain was ranked fourth by the adjusted DRC and API statistics. The ranks for the remaining domains varied across importance measures. For example, the SF-36 vitality domain resulted in a rank of 6 using the API and a rank of 10 when the SLRC was used.

4. Discussion and Conclusions

Relative importance measures have been described in social science disciplines [49-51], but there have been little written about them in the health sciences literature. These measures have potential benefit in HRQOL research for describing group differences on multiple domains.

Previous research has shown that importance measures do not always result in consistent rankings [12, 33]. Dissimilarities in rankings may arise, in part, because of the data characteristics and the assumptions that underlie the models on which relative importance measures are based [14]. This was evident in the numeric example, where two disease-specific domains were consistently ranked as the most important by all but two measures. However, for the other domains, there were differences in the ranks. The presence of suppressor and/or collinear variables may have contributed to these differences. Also, given that the procedures are not equally sensitive to non-normality and covariance homogeneity, these data characteristics may have contributed to the differences in rankings.

We recommend that researchers undertake a careful descriptive assessment of their data to assess the tenability of model assumptions before choosing a relative importance measure. While it is possible to conduct statistical tests of departures from the assumptions of normality and

variance heterogeneity [52-54], these tests have some limitations. Tests of variance homogeneity are known to be sensitive to departures from a normal distribution [51, 55], and tests of normality are sensitive to sample size [55]. Therefore, descriptive measures of skewness and kurtosis and ratios of group variances should be preferred.

An internal or external validation of the ranks should also be considered, in order to assess the generalizability of the results. For example, a split-sample validation might be conducted. Resampling-based methods such as the bootstrap have also been proposed to describe the sampling variability in the ranks [56-57].

There are additional considerations when conducting a relative importance analysis. The conclusion that one HRQOL domain is more important than another can only be applied to the set of domains under investigation. Hence, changing the domains included in the analysis may result in different conclusions about relative importance. Importance may not correspond with clinical significance [58-59]. As well, relative importance may be associated with covariates such as age and sex. Stratified analyses might be conducted to assess the influence of covariates on the results. Alternatively, covariates can be incorporated into the LR model and by adopting a multivariate analysis of covariance model [60].

Relative importance measures have a number of potential uses for researchers who conduct studies about HRQOL. They can be used to develop parsimonious statistical models. In exploratory research, they can be used to identify a small set of domains on which to focus in future studies. The measures could be used to assign weights to the domains when using a multiple testing procedure to control the overall probability of a Type I error; procedures in which the weights are assigned *a priori* have been shown to result in substantially improved power to detect group differences on the most important domains [61-62]. The measures have

other practical applications. For example, when comparing chronic disease patients to healthy controls, relative importance measures can provide information about the domains on which the disease has the greatest effect, which might be useful in the development of clinical interventions.

Acknowledgements

This research was supported by a Canadian Institutes of Health Research (CIHR) Vanier Graduate Scholarship to the first author, funding from the Manitoba Health Research Council and a CIHR New Investigator Award to the second author, funding from a Crohn's and Colitis Foundation of Canada Research Investigator Award and the Bingham Chair in Gastroenterology to the last author and funding from a CIHR Operating Grant to the research team.

Conflicts

Dr. Lix has received funding from Amgen in the form of an unrestricted research grant. In the past year Dr. Bernstein has received consulting fees from Abbott Canada and an unrestricted educational grant from Axcan Pharma.

References

1. Fairclough, D. L. (2002). *Design and Analysis of Quality of Life Studies in Clinical Trials*. New York: Chapman & Hall.
2. Sheehan-Holt, J. K. (1998). MANOVA simultaneous test procedures: The power and robustness of restricted multivariate contrasts. *Journal of Educational and Psychological Measurement, 58*, 861 – 881.
3. Wu, Y., Hu, W., Xia, Y., Ma, J., Liu, M., & Cui, N. (2006). Quality of life of nasopharyngeal carcinoma survivors in Mainland China. *Quality of Life Research, 16*, 65 – 74.
4. Laaksonen, C., Aromaa, M., Heinonen, O. J., Koivusilta, L., Koski, P., Suominen, S., Vahlberg, T., & Salantera, S. (2008). Health related quality of life in 10-year-old school children. *Quality of Life Research, 17*, 1049 – 1054.
5. Dunn, W. S., Mount, M. K., Barrick, M. R., & Ones, D. S. (1995). Relative importance of personality and general mentality ability in managers' judgments of applicant qualifications. *Journal of Applied Psychology, 80*, 500 – 509.
6. Hobson, C. J., & Gibson, F. W. (1983). Policy capturing as an approach to understanding and improving performance appraisal: A review of the literature. *Academy of Management Review, 8*, 640 – 649.
7. Healy, M. J. R. (1990). Measuring importance. *Statistics in Medicine, 9*, 633 – 637.
8. Curran, D., Molenberghs, G., Fayers, P. M., & Machin, D. (1998). Incomplete quality of life data in randomized trials: missing forms. *Statistics in Medicine, 17*, 697 – 709.
9. Rose, M. S., & Koshman, M. L. (1999). Statistical issues encountered in the comparison of health-related quality of life in diseased patients to published general population norms: Problems and solutions. *Journal of Clinical Epidemiology, 52*, 405 – 412.

10. Beaumont, J. L., Lix, L. M., Yost, K. J., & Hahn, E. A. (2006). Application of robust statistical methods for sensitivity analysis of health-related quality of life outcomes. *Quality of Life Research, 15*, 349 – 356.

11. Menard, S. (1995). *Applied Logistic Regression Analysis*. Thousand Oaks, CA: Sage.

12. Thomas, D. R., Zumbo, B. D., Zhu, P, & Dutta, S. (2008). On measuring the relative importance of explanatory variables in a logistic regression. *Journal of Modern Applied Statistical Methods, 7*, 21 – 38.

13. Traxel, N. M., & Azen, R. (2006). *Predictor importance in logistic regression: An extension of dominance analysis*. Poster presented at the 114th Annual Convention of the American Psychological Association, New Orleans, LA.

14. Tonidandale, S., & LeBreton, J. M. (2010). Determining the relative importance of predictors in logistic regression: An extension of relative weight analysis. *Organizational Research Methods, 13,* 767 – 781.

15. Huberty, C. J., & Wisenbaker, J. M. (1992). Variable importance in multivariate group comparisons. *Journal of Educational Statistics, 17*, 75 – 91.

16. Thomas, D. R. (1992). Interpreting discriminant functions: a data analytic approach. *Multivariate Behavioral Research, 27*, 335 – 362.

17. Laaksonen, C., Aromaaa, M., Heinonen O. J. et al. (2008). Health related quality of life in 10-year old school children. *Quality of Life Research, 17*, 1049 – 1054.

18. Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology, 65*, 145 – 153.

19. Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Educational Psychology, 25*, 241 – 286.

20. Agresti, A. (1996). *An Introduction to Categorical Data Analysis*, New York: Wiley.

21. Menard, S. (2004). Six approaches to calculating standardized logistic regression coefficients. *The American Statistician, 58*, 218 – 223.

22. Agresti, A., & Finlay, B. (1986). *Statistical Methods for the Social Sciences, 2nd Ed.* San Francisco: Dellen/Macmillan.

23. Kaufman, R. L. (1996). Comparing effects in dichotomous logistic regression: A variety of standardized coefficients. *Social Science Quarterly, 77*, 90 – 110.

24. Nagelkerke, N. J. D. (1991). A note on a general definition of coefficient of determination. *Biometrika, 78*, 691 – 692.

25. Cox, D. R., & Snell, E. J. (1989). *The analysis of binary data (2nd ed.)*. London: Chapman and Hall.

26. Estrella, A. (1998). A new measure of fit for equations with dichotomous dependent variables. *Journal of Business and Economic Statistics, 16*, 198 – 205.

27. Pratt, J. W. (1987). Dividing the indivisible: Using simple symmetry to partition variance explained. In T. Pukkila and S. Puntanen (Eds), *Proceedings of the Second International Conference in Statistics* (p. 245 – 260). Tampere, Finland: University of Tampere.

28. Thomas, D. R., Hughes, E., & Zumbo, B. D. (1998). On variable importance in linear regression. *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement, 45*, 253 – 275.

29. Ochieng, C. O., & Zumbo, B. D. (2001). *Examination of a variable ordering index in linear regression models: An assessment of the relative Pratt index in Likert data*. Paper presented at the Bob Conry Conference on Measurement Evaluation and Research and Methodology,Vancouver, Canada.

30. Witherill, G. G. (1986). *Regression Analysis with Applications*. London: Chapman & Hall.

31. Azen, R., & Budescu, D. V. (2003). Comparing predictors in multivariate regression models: An extension of dominance analysis. Journal *of Educational and Behavioral Statistics, 31,* 157 – 180.

32. Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research, 35*, 1 – 19.

33. Thomas, D. R., & Zumbo, B. D. (1996). Using a measure of variable importance to investigate the standardization of discriminant coefficients. *Journal of Educational and Behavioral Statistics, 21*, 110 – 130.

34. Huberty, C. J., & Olejnik, S. (2006). *Applied MANOVA and Discriminant Analysis*. New York: Wiley.

35. Pohar, M., Blas, M., & Turk, S. (2004). Comparison of logistic regression and linear discriminant analysis: A simulation study. *Metodoloski zveski, 1*, 143 – 161.

36. Jennrich, R. I. (1977). Stepwise discriminant analysis. In Enslein, K, Ralston. A, & Wilf, H. S. (Eds.). *Mathematical Methods for Digital Computers, Vol 3*. John Wiley & Sons: New York.

37. Rencher, A. C. (1992). Interpretation of canonical discriminant functions, canonical variates, and principal components. *The American Statistician, 46*, 217 – 225.

38. Finch, W. H., & Laking, T. (2008). Evaluation of the use of standardized weights for interpreting results from a descriptive discriminant analysis. *Multiple Linear Regression Viewpoints, 34*, 19 – 34.

39. Gibbons, L. B., & Hosmer, D. W. (1991). Conditional logistic regression with missing data. *Communications in Statistics, Part B, 20*, 109 – 120.

40. Little, R. J. A., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data, 2nd Ed*. New Jersey: Wiley.

41. SAS Institute Inc. (2008). *SAS/STAT User's Guide, Version 9.2.* Cary, NC: SAS Institute Inc.

42. Lix, L. M., Graff, L. A., Walker, J. R, L., et al. (2008). Longitudinal study of quality of life and psychological functioning for active, fluctuating, and inactive disease patterns. *Inflammatory Bowel Diseases, 14*, 1575 – 1584.

43. Graff, L. A., Walker, J. R., Lix, L. M., et al. (2006). The relationship of inflammatory bowel disease type and activity to psychological functioning and quality of life. *Clinical Gastoroenterology and Hepatology, 4*, 1491 – 1501.

44. Clara, I., Lix, L. M., Walker, J. R., et al. (2009). The Manitoba IBD index: Evidence for a new and simple indicator of IBD activity. *American Journal of Gastroenterology, 104*, 1754 – 1763.

45. Guyatt, G. H., Mitchell, A., Irvine, E. J., et al. (1989). A new measure of health status for clinical trials in inflammatory bowel disease. *Gastroenterology, 96*, 804 – 810.

46. Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care, 30*, 473 – 483.

47. Irvine, E. J. (1996). Effects of budesonide therapy on quality of life in active Crohn´s disease. *Respiratory Clinical Forums, 18*, 81 – 89.

48. Shapiro, S. S.; Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika 52,* 591– 611.

49. LeBreton, J. M, Tonidandel, S. (2008). Multivariate relative importance: Extending relative weight analysis to ultivariate criterion spaces. *Journal of Applied Psychology, 93*, 329-345.

50. Lievens, F., Van Hoye, G., & Schreurs, B. (2005). Examining the relationship between

employer knowledge dimensions and organizational attractiveness: An application in a military context. *Journal of Occupational and Organizational Psychology, 78*, 553 – 572.

51. Glomb, T. M., Kammeyer-Mueller, J. D., & Rotundo, M. (2004). Emotional labor demands and compensating wage differentials. *Journal of Applied Psychology, 84*, 700 – 714.
52. Hawkins, D. M. (1981). A new test for multivariate normality and homoscedasticity. *Technometrics, 23*, 105 – 110.
53. Box, G.E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika, 36*, 317 – 346.
54. Mardia, K. V. (1980). Tests of univariate and multivariate normality. In P.R. Krishnaiah (Ed.), *Handbook of Statistics, Vol 1* (pp 279 – 320). Amsterdam: North-Holland.
55. Shapiro, S. S., Wilk, M. B., & Chen, H. J. (1968). A comparative study of various tests for normality. *Journal of the American Statistical Association, 63,* 1343 – 1372.
56. Huberty, C. J. (1975). The stability of three indices of relative variable contribution in discriminant analysis. *Journal of Experimental Education, 44*, 59 – 64.
57. Dalgleish, L. I. (1994). Discriminant analysis: Statistical inference using the jackknife and bootstrap procedures. *Psychological Bulletin, 116*, 498 – 508.
58. Guyatt, G. H., Osoba, D., Wu, A. W., Wyrwich, K. W., Norman, G. R., & the Clinical Significance Consensus Meeting Group. (2002). Methods to explain the clinical significance of health status measures. *Mayo Clinic Proceedings, 77*, 371 – 383.
59. Wyrwich, K. W., Nienaber, N. A., Tierney, W. M., & Wolinsky, F. D. (1999). Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Medical Care, 37*, 469 – 478.
60. Kim, K., & Timm, N. (2007). *Univariate and Multivariate General Linear Models: Theory*

and Applications with SAS, 2nd ed. New York: Chapman & Hall/CRC.

61. Benjamini, Y., & Hochberg, Y. (1997). Multiple hypotheses testing with weights. *Scandinavian Journal of Statistics, 24*, 407 – 418.

62. Maurer, W., Hothorn, L. A., & Lechmacher, W. (1995). Multiple comparisons in drug clinical trials and preclinical assays: A priori ordered hypotheses. In Vollman, J., (Ed), *Biometrie in der chemische-pharmazeutichen Industrie* (Vol. 6). Stuttgart: Fischer Verlag.

Table 1. Characteristics of Manitoba IBD Cohort Study Participants

	Crohn's Disease (*n* = 187)	**Ulcerative Colitis** (*n* = 169)
Active disease (%)	74	66
Mean age, years (SD)	38.5 (14.6)	43.0 (14.7)
Female (%)	61	58
Marital Status		
Married or common-law	64	71
Single, never married	28	17
Other	8	12
Education		
No post secondary	43	35
Trade school, diploma	32	34
University	25	31
Ethnicity, Caucasian	93	88
Mean disease duration, years (SD)	4.4 (2.1)	4.3 (2.1)

Note: SD = standard deviation

Table 2. Descriptive Statistics for HRQOL Domain Scores

	Active Disease (n_1 = 244)	**Inactive Disease** (n_2 = 105)
IBDQ		
Bowel Symptoms	4.92 (1.03)	6.08 (0.76)
Emotional Health	4.81 (1.05)	5.85 (0.89)
Social Function	4.09 (1.18)	5.19 (1.05)
Systemic Symptoms	5.62 (1.35)	6.65 (0.64)
SF-36		
Bodily Pain	60.78 (24.15)	77.45 (26.11)
Role Physical	63.48 (29.07)	83.65 (24.08)
General Health	43.40 (19.52)	59.18 (17.01)
Mental Health	60.33 (14.11)	66.62 (12.47)
Physical Functioning	77.49 (21.73)	91.11 (14.41)
Role Emotional	76.06 (23.98)	85.82 (20.11)
Social Functioning	63.74 (27.20)	78.85 (27.10)
Vitality	46.13 (16.39)	57.84 (14.49)

Note: Values reported are mean (SD); IBDQ = Inflammatory Bowel Disease Questionnaire; SF-36 = 36-item Short Form Questionnaire

Table 3. Significance Test Results and Numeric Values of Relative Importance Measures for HRQOL Domains

Domain	*t*-statistic	SLRC	PI	API	RRW	SDFC	DRC	ADRC	FTR
IBDQ									
Bowel Symptoms	10.430*	0.463	0.471	0.376	0.006	0.587	0.542	0.44	5.034
Emotional Health	8.840*	0.309	0.28	0.223	0.088	0.428	0.347	0.282	4.033
Social Function	7.500*	0.183	0.165	0.132	0.050	0.044	-0.031	0.000	5.072
Systemic Symptoms	7.980*	0.145	-0.117	-	0.083	0.083	-0.062	0.000	14.334
SF-36									
Bodily Pain	5.690*	0.103	0.066	0.053	0.030	0.103	0.057	0.047	0.504
Role Physical	6.220*	0.015	-0.010	0.000	0.369	0.037	-0.022	0.000	6.099
General Health	6.930*	0.135	0.095	0.076	0.002	0.226	0.149	0.l21	12.334
Mental Health	3.790*	0.143	-0.059	-	0.008	0.1910	-0.072	-	0.952
Physical Functioning	5.890*	0.169	0.113	0.090	0.106	0.185	0.106	0.086	8.329
Role Emotional	3.640*	0.171	-0.066	-	0.004	0.120	-0.043	-	0.508
Social Functioning	4.770*	0.026	0.015	0.012	0.180	0.027	0.013	0.011	0.011
Vitality	6.080*	0.074	0.049	0.039	0.076	0.029	0.017	0.014	6.911

Note: IBDQ = Inflammatory Bowel Disease Questionnaire; SF-36 = 36-item Short Form Questionnaire; * denotes a test statistic that is significant at $\alpha = .004$; SLRC = standardized logistic regression coefficient; PI = Pratt Index; API = adjusted Pratt Index; RRW = Rescaled relative weight; SDFC = standardized discriminant function coefficient; DRC = Discriminant ratio coefficient; ADRC = Adjusted discriminant ratio coefficient; FTR = *F*-to-remove statistic. Each measure is reported as an absolute value, except PI and DRC; - indicates a potential suppressor variable that was excluded from the analysis.

Table 4. Rank Order of HRQOL Domains based on Relative Importance Measures

Domain	SLRC	API	RRW	SDFC	ADRC	FTR
IBDQ						
Bowel Symptoms	1	1	10	1	1	7
Emotional Health	2	2	4	2	2	8
Social Function	3	3	7	9	9	6
Systemic Symptoms	6	-	5	8	9	1
SF-36						
Bodily Pain	9	6	8	7	5	11
Role Physical	12	9	1	10	9	5
General Health	8	5	12	3	3	2
Mental Health	7	-	9	4	-	9
Physical Functioning	5	4	3	5	4	3
Role Emotional	4	-	11	6	-	10
Social Functioning	11	8	2	12	7	12
Vitality	10	7	6	11	6	4

Note: 1 = highest relative importance; IBDQ = Inflammatory Bowel Disease Questionnaire; SF-36 = 36-item Short Form Questionnaire; SLRC = standardized logistic regression coefficient; API = adjusted Pratt Index; RRW = rescaled relative weight; SDFC = standardized discriminant function coefficient; ADRC = adjusted discriminant ratio coefficient; FTR = *F*-to-remove; - indicates a potential suppressor variable that was excluded from the analysis.

Appendix A. Additional Formulae used in Calculating Measures of Relative Importance

Measures Based on the LR Model

The estimated LR coefficient for the *k*th domain ($k = 1, \ldots, m$) can be written as

$$\hat{\beta}_k = \frac{r_{\text{logit}(\hat{p})k} - R^2_{(-k)} R^2_{k|(-k)}}{1 - R^2_{k|(-k)}}, \tag{A-1}$$

where $r_{\text{logit}(\hat{p})k}$ is the correlation between the *k*th domain and the logit of the predicted probabilities, $R^2_{(-k)}$ is the R^2 value for a LR model in which the *k*th domain is excluded, and $R^2_{k|(-k)}$ is the R^2 value for a model in which the *k*th domain is regressed on the other $(m - 1)$ domains.

Pratt's index can also be expressed as,

$$d_k = \frac{\hat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{X}^{\mathrm{T}} \mathbf{Q} \mathbf{X}_{[k]} \hat{\beta}_k}{\hat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{X}^{\mathrm{T}} \mathbf{Q} \mathbf{X} \hat{\boldsymbol{\beta}}}, \tag{A-2}$$

where $\mathbf{X}$ is the $N \times m$ data matrix, $\mathbf{X}_{[k]}$ is the $N \times 1$ vector of measurements on the *k*th domain, $\hat{\boldsymbol{\beta}}$ is a $m \times 1$ vector of estimated unstandardized LR coefficients, $\mathbf{Q} = \mathbf{I}_N - (\mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} / N)$, $\mathbf{I}_N$ is a $N \times N$ identity matrix, $\mathbf{1}_N$ is a $N \times 1$ matrix of ones, and $^{\mathrm{T}}$ is the transpose operator.

Measures Based on the MANOVA Model

The vectors of discriminant function coefficients corresponds to the eigenvectors associated with $\mathbf{E}^{-1}\mathbf{H}$, where

$$\mathbf{E} = \sum_{j=1}^{2} \sum_{i=1}^{n_j} (\mathbf{X}_{ij} - \overline{\mathbf{X}}_j)(\mathbf{X}_{ij} - \overline{\mathbf{X}}_j)^{\mathrm{T}}, \tag{A-3}$$

is the error sum of squares and cross product matrix, and

$$\mathbf{H} = \sum_{j=1}^{2} n_j (\mathbf{X}_j - \overline{\mathbf{X}})(\mathbf{X}_j - \overline{\mathbf{X}})^{\mathrm{T}}, \tag{A-4}$$

is the hypothesis sum of squares and cross product matrix. The number of statistically significant discriminant functions is $c = \min(m, g - 1)$. The discriminant function score, z_{ij}, for the ith study participant in the jth ($i = 1,\ldots, n_j; j = 1, 2$) group is,

$$z_{ij} = \hat{\mathbf{a}}\mathbf{X}_{ij}. \tag{A-5}$$

The discriminant function coefficient for the kth variable can also be expressed as

$$\hat{a}_k = -\log(\frac{n_2}{n_1}) - \frac{1}{2}(\overline{\mathbf{X}}_1 + \overline{\mathbf{X}}_2)^{\mathrm{T}}\mathbf{S}^{-1}(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) + u_k \tag{A-6}$$

where u_k is the kth element of $\mathbf{S}^{-1}(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2)$, $\mathbf{S}$ is the pooled sample covariance matrix, and $\overline{\mathbf{X}}_j$ is the vector of means for the jth group.

An equivalent formula for computing the F-to-remove statistic for the kth domain is

$$F_{(k)} = \frac{k_2(\hat{a}_k / s_{(kk)})^2}{(\bar{z}_1 - \bar{z}_2) + k_3 - (\hat{a}_k / s_{(kk)})^2}. \tag{A-7}$$

where $k_2 = (n_1 + n_2 - 2 - m)$, $k_3 = (n_1 + n_2)(n_1 + n_2)/n_1 n_2$, $\hat{a}_k$ is the discriminant function coefficient for the kth domain, $\bar{z}_1$ and $\bar{z}_2$ are the group means for the discriminant function score corresponding to $\hat{\mathbf{a}}$, and $s_{(kk)}$ is the positive square root of the kth diagonal element of the inverse of $\mathbf{E}$, the error sums of square and cross product matrix.

Appendix B. Correlations among HRQOL Domains for Active and Inactive Disease Groups

	IBDQ BS	IBDQ EH	IBDQ SF	IBDQ SS	SF36 BP	SF-36 RP	SF-36 GH	SF-36 MH	SF-36 PF	SF-36 RE	SF-36 SF	SF-36 VT
IBDQ BS	1	0.63	0.57	0.65	0.43	-0.43	0.41	0.29	0.27	-0.28	0.36	0.42
IBDQ EH	0.70	1	0.71	0.71	0.38	-0.53	0.53	0.69	0.34	-0.57	0.49	0.59
IBDQ SF	0.57	0.72	1	0.60	0.43	-0.64	0.43	0.39	0.47	-0.42	0.57	0.43
IBDQ SS	0.51	0.66	0.57	1	0.43	-0.55	0.43	0.46	0.45	-0.40	0.43	0.59
SF36 BP	0.27	0.25	0.37	0.39	1	-0.66	0.38	0.29	-0.41	-0.40	0.62	0.39
SF36 RP	0.43	0.56	0.62	0.44	-0.69	1	-0.48	-0.44	-0.60	-0.58	-0.64	-0.52
SF36 GH	0.19	0.27	0.27	0.40	0.48	-0.42	1	0.49	-0.48	-0.35	0.41	0.63
SF36 MH	0.37	0.59	0.44	0.44	0.52	-0.61	0.38	1	0.21	-0.67	0.49	0.60
SF36 PF	0.25	0.38	0.31	0.45	0.25	-0.44	0.30	0.19	1	-0.35	0.37	0.41
SF36 RE	-0.37	-0.55	-0.55	-0.40	-0.55	0.71	-0.45	-0.68	-0.36	1	-0.54	-0.50
SF-36 SF	0.17	0.26	0.36	-0.30	0.80	-0.66	0.49	0.50	0.23	-0.58	1	0.45
SF-36 VT	0.32	0.59	0.42	0.61	0.41	-0.47	0.48	0.57	0.33	-0.61	0.51	1

Note: Correlations for the active disease group are on the upper diagonal and the correlations for the inactive disease group are on the lower diagonal; IBDQ = Inflammatory Bowel Disease Questionnaire; SF-36 = 36-item Short Form Questionnaire; BS = bowel symptoms; EH = emotional health; SF = social functioning; SS = systemic symptoms; BP = bodily pain; RP = role physical; GH = general health; MH = mental health; PF = physical functioning; RE = role emotional; VT = vitality.