WORK AND WELLNESS INDEX

Lundholm I. [1], **Lundholm H.** [1]
FHB Technology, Snövägen 68b, 976 34 Luleå, Sweden
e-mail: ilundholm@bredband.net

Integration of data is the foundation for an optimal health system. The principal aim is to motivate individuals to make improvements in their personal health and lifestyle. The management needs essential information in order to evaluate outcome and to guide corporate wellness programs.

Riskmanagement, multivariate, models, health

1 Introduction

Systematic work environment management includes measuring psychological and social conditions and the employees' general health. The results should include assessment of risks and appropriate action plans. In Sweden this is outlined in SAM, a documentation supervised by the Swedish work environment authory (2006).

2 Objectives

The primary objective of this study was to use multivariate data analysis (MDA) to assist in powerful health risk assessments on a regular basis. The resulting multivariate models were interpretated and compared with published and validated information, associated with the metabolic syndrome and with mental stress in the environment.

3 Methods

The participants in the study were men and women in the age of 20 to 65 years, living in Norrbotten, Sweden. Individuals included in the global statistical models have different backgrounds and are carefully selected during a period of several years to resemble a normal distributed cohort of both men and women and of resonable size.

3.1 Health assessment system

The work and wellness health system is built on a multi-level strategy. This includes a questionnarie addressing 8 major factors with a grand total of 110 questions, efficient clinical procedures and a broad fitness test. The lab and fitness results are presented graphically and used in personal communications with the participants. When obvious health deviations are detected, a visit to a doctor is recommended for follow-up. The assessment system securely gathers and records results for each individual. This ensures that time dependent factors in health and life style can be calculated and presented. For employee groups brief statistical patterns and general health can be summarized in a company report. Statistical methods in general, see Hjorth (1998).

3.2 Multivariate data analysis

Large data tables with many variables and a limited amount of observations (objects) require the need of multivariate projection methods, like principal component analysis (PCA). Some of the advantages are that it separates regularities from noise, copes with missing data and provides informative diagnostic and graphical tools. The theoretical background and principles for PCA and other important techniques are outlined by Wold et al. (2001).

A PCA-model of a data matrix **X** is represented by calculating a number of pricipal components of a training set. Typically, 2 to 5 principal components are sufficient to approximate a data table. The number of components is in practise optimised by using cross-validation described by Wold (1978).

By using PCA a data table **X** is modelled as $\mathbf{X} = 1*x' + \mathbf{T}*\mathbf{P'} + \mathbf{E}$

The first term 1*x', represents the variable averages and originates from the pre-processing step. The second term, the matrix product **T***P', models the structure, and the third term, the residual matrix **E**, contains the noise.
In the graphical interpretation, the two low-dimensional matrices, **T** an **P**', can be projected into score and loading plots in one, two or three dimensions. The meanings of the scores are given by the loadings. The loadings define the orientation of the PC plane with respect to the original X-variables. It also give information of the *magnitude* (large or small correlation) and the *manner* (positive or negative correlation) in which the measured variables contributes to the scores. An important concept is the transformation of a combination of measured variables into *latent variables*. A latent variable can be seen as a property in the material to be interpretated.

4 Results

The study has resulted into several separate multivariate model compositions. A primary model includes clinical and physiological fitness data together with some basic lifestyle information. It is found necessary to establish separate models for men and women but still use the same variable composition.The health assessment questionnarie includes 8 indexgroups. Extensive data analysis has shown a limited correllation structure betweeen the latter range of data and the clinical/fitness data assembly. A secondary independent model, based on the questionnarie, visualize the ability for an individual to handle tensions and stress responses, Antonovsky (1991).

4.1 The metabolic syndrome

The metabolic syndrome, also called the insuline resistence syndrome, represents a group of risk factors for cardiovascular disease. A combination of an increased waist circumference due to accumulated visceral fat, dyslipidemia characterized of high levels of triglycerides and a low HDL-cholesterole together with hypertoni (high blood preassure) means a strong increase in the risk for serious illness or sudden death due to artheroscleros or heart failure, Ekman, Arnetz p.160-171 (2005). Traditionally

prevention has focused on single risk factors such as high blood pressure, high lipid levels or diabetes. Now there is an aim at thinking of the problem as corresponding to a multifactorial cause, where single risk factors tend to support each other, Rosengren (2006). A range of studies shows an inverse connection between the level of physical activity or physical capacity (condition) and cardiovascular diseases and mortality. The strong protecting effect from physical activity is explained by multiple and complex effects on lipoproteins and lipidmetabolism, balance in hormone levels, blood pressure and sensitivity to insulin response to name a few, Hellénius (2006). A problem with insuline resistence as such, when the biological response to insuline is insufficient, is that it is difficult to establish this state by clinical methods. 'The risk of congestive heart failure (CHF) increases already in the sub-clinical phase of impaired glucose regulation that foregoes clinical manifest diabetes'. Ingelsson, p 43 (2005).

In dataanalysis with PCA, applicated on a selected dataset, we obtain a 4-dimensional PC-model with an explained variation of 53 %. The scores of the PC are accompanied by the corresponding loadings. A scatter plot of the loadings of the first component versus the loadings of the third component (p_1 vs. p_3) is presented in figure 1.

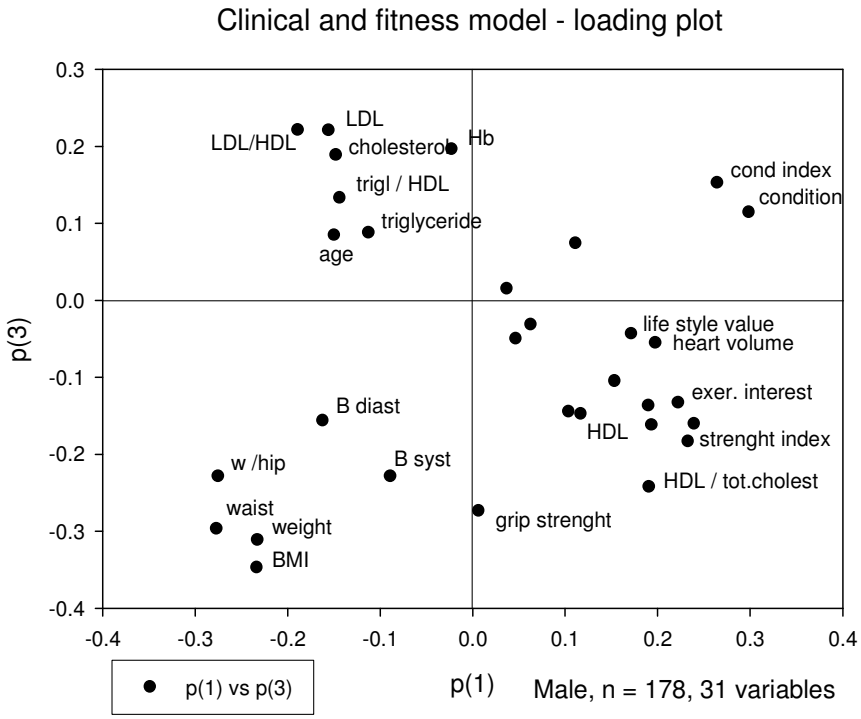

Figure 1. PCA loading plot of the loadings of the first and third principal components of the male Clinical/Fitness dataset. In the plot, several small clusters of risk factors to the left side and protective factors to the right side are visible.

In the loading plot, variables with a protecting ability are recognized to the right and the variables associated with the metabolic syndrome, to the left. The latter group consists of two distinct subgroups; the lower one is composed of body weight, BMI, waist circumference and the waist/hip ratio. The upper subgroup (still to the left) is a clustre of the lipid family of variables. High density lipoprotein (HDL), however, is located in the lower right quadrant. A high level of HDL is a protective factor, Rosengren (2006). The variables representing blood-pressure (diastolic and systolic) are found to the lower, left. In some respect, high blood-pressure values are associated with high body weight and low condition and strength values, Hellénius (2006).

Each individual, included in the model, has a corresponding score value in this projection of the multivariate PC space. With a high positive score value in the first component (t_1), a good combination of health indicators can be expected. Subsequently the opposite holds for a person with a negative score value. In a plot of a single variable such as the physical condition (ml oxygen/ kg*min) versus the individual score values for the first component, a distinct trend in data can be recognized. The sig-saw pattern in the connecting line between the points reflects that 30 other variables influence the score value. Condition values in this upper range can be obtained by a combination of physical exercise and a healthy lifestyle, Nilsson (1998).
In figure 2, a simple regression of the point values for male condition, reveals that more than 40 (ml O_2/kg*min) corresponds to a positive score value in the first component of the fitness model. In a similar plot with t (1) vs. waist circumference, a similar strong but inverse correllation is revealed. (Waist 93 cm at t (1) = 0, not shown in the text).

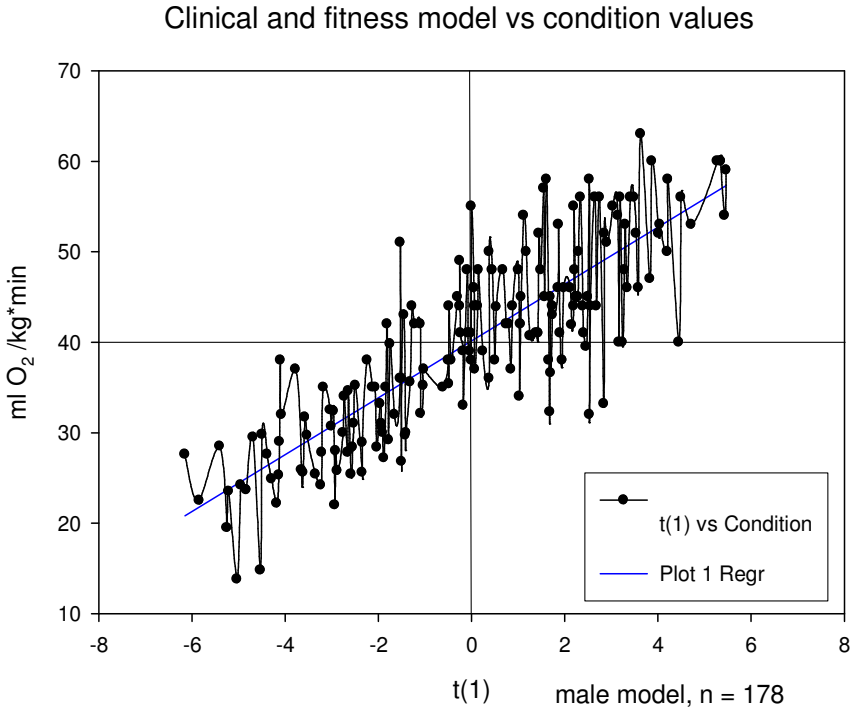

Figure 2. A male fitness model with the individual score values compaired with a corresponding single variable - physical condition.

The composition of variables in the loading plot, expressed as the *latent variable* or the quality represented by the principal components could also be explained as the balance between catabolic and anabolic processes occuring in the human body. Anabolic processes correspond to increased levels of sex steroids and growth hormone and bring a reduced activity in the HPA (the *hypothalamic-pituitary-adrenocortical* system) and SAM (the *sympathetic-adreno-medullary* system), Ekman, Arnetz p.372-91 (2005).

4.2 Resistence to psychosocial stress

Mental and psychosocial stress can be defined as a changing game between the demands from the surrounding world and the possibilities for an individual to cope with these demands. Activation of cortisol release and circuit control (HPA system) is both age dependent and multifactorial in nature. Severe body pain and sleep disorder are common features that is incorporated in disorder states, Lundberg, Wentz (2004).

In a multivariate, unisex, psychosocial stress model, a clear and strong correllation occurs for the KASAM, the psychosocial work environment, the bodypain and the sleep/recovery variable groups. Low score values in the first component of the model tend to predict a presence of stress disorder. In figure 3, an example of a common variable correllation structure hidden in such a datamatrix is plotted. Balance in strain and resources are well described by Perski (2002).

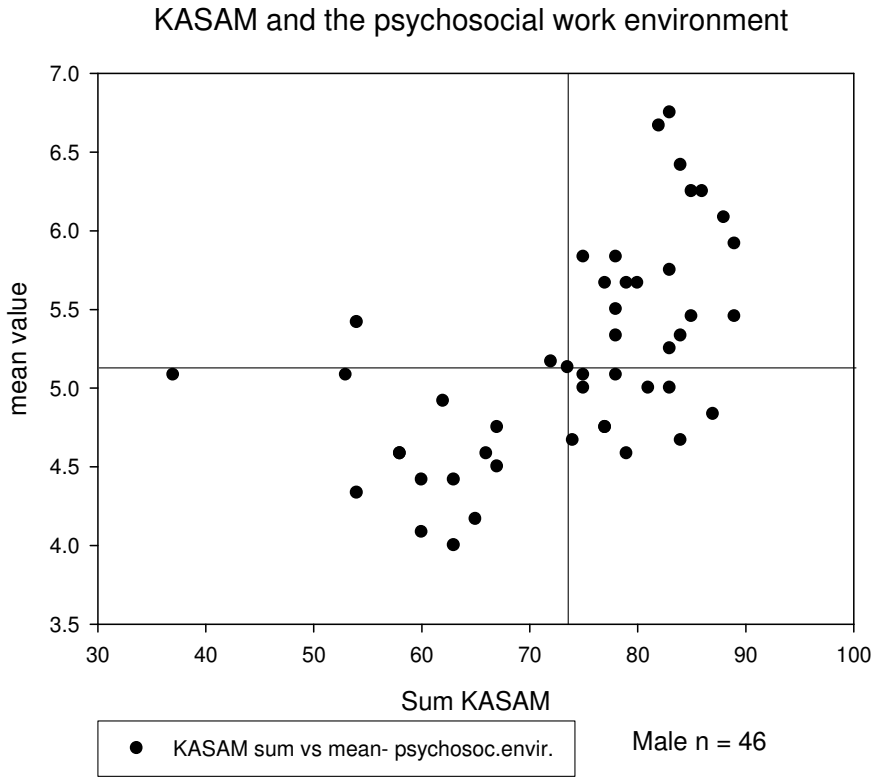

Figure 3. A strong individual KASAM tends to increase the possibility to more frequently mark a positive alternative in a questionnary regarding the psychosocial work environment. The dataset is part of a summary for a single male employee group with similar occupations.

5 Discussion

In accordance with the current view at the Medical Products Agency, multiple risk factors for major public health problems are joined together in order to more effectivly assess the total risk for future diseases. At a company level, by taking into account protective factors and considering the time dependent dimension, reliable information can be presented to each individual. Leaders on different levels in a company can recieve a summary of important general information in line with the requirements outlined in SAM. The multivariate approach enables clients in due time to halt and reverse negative trends and weaknesses, that often are difficult to identify with standard methods, and thus follow a new healthy lifepath. This requires the support from a regular company schedule with a clear strategy for achievements in human resources.

Health models can be extended to monitor, not only physical and mental individual patterns, but also group dynamics in the working environment. This includes attitudes, questions concerning management and the overall innovative and creative climate in a company, Ekvall (2006). Score values on group level from such models gives perspective on the competing ability and the productivity in different company units.

6 References

Antonovsky, A. 1991. Hälsans mysterium. Natur och Kultur, Stockholm.
Ekman, R., & Arnetz, B. 2005. Stress, Molekylerna – individen – organisationen – samhället. Liber, Stockholm.
Ekvall, G. 2006. Psykologtidningen nr 13. Sveriges Psykologförbund.
Eriksson, L., E. Johansson, Kettaneh-Wold, N., Wold, S. Multi- and Megavariate Data Analysis. Principles and Applications. 2001. Umetrics Academy, Umeå.
Hellénius, M-L. 2006. Livsstilsåtgärder – fysisk aktivitet. Information från Läkemedelsverket 3:2006. Läkemedelsverket.
Hjorth, U. Statistisk slutledning i ekonomi och teknik. Studentlitteratur, Lund.
Ingelsson, E. 2005. Insulin Resistence and Inflammation as Risk Factors for Congestive Heart Failure. Acta Universitatis Upsaliensis 2005, Uppsala.
Lundberg, U., Wentz, G. 2004. Stressad hjärna, stressad kropp. Wahlström & Widstrand, Stockholm.
Nilsson, J. 1998. Puls- och laktatbaserad träning. SISU Idrottsböcker, Farsta.
Perski, A. 2002. Ur balans. Bonniers, Stockholm.
Rosengren, A. 2006. Epidemiologi – riskfaktorer. Information från Läkemedelsverket 3:2006. Läkemedelsverket.
Wold, S. 1978. Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models, Technometrics, 20, 397-405.