MULTIVARIATE REGRESSION MODEL OF IMPEDANCE OF NORMAL AND CHEMICALLY IRRITATED SKIN SHOWS PREDICTIVE ABILITY

P. Åberg[1], I. Nicander[2], P. Geladi[3], S. Ollmar[1]
[1] Medical Engineering, Karolinska Institutet, Huddinge, Sweden
[2] Dermatology, Huddinge University Hospital, Huddinge, Sweden
[3] Dept. of Chemistry, Umeå University, Umeå, Sweden

Abstract - **We present predictive models that can foresee how skin will react when exposed to chemicals. Skin impedance spectra, 31 frequencies between 1 and 1000 kHz at five depth settings, were collected before and after application of chemicals on volar forearms of volunteers. Tegobetaine and sodium lauryl sulphate were used to induce the irritations. The spectra were filtered using orthogonal signal correction (OSC). The relation between skin impedance of normal and chemically irritated skin was modelled using partial least squares regression (PLS). The predictive ability of this model is demonstrated for two irritants, and additional studies are required to establish this property for other chemicals.**

Keywords - **Skin impedance, skin irritation, multivariate analysis, partial least squares regression, PLS, orthogonal signal correction, OSC**

I. INTRODUCTION

Bio-impedance detection is a non-invasive, rapid and reliable method in characterising the properties of skin and other biological tissues. It has previously been shown that skin impedance is well suited in monitoring skin irritations. The purpose of this study is to compare and correlate impedance of two conditions of skin; unaffected normal skin and skin irritated by chemicals, using multivariate projection methods.

II. METHODOLOGY

A. Clinical

Locations on the forearms of 21 healthy volunteers were exposed to two surfactant solutions, tegobetaine (TEG) and sodium lauryl sulphate (SLS), for 24 hours using Finn chambers. The concentrations were 1% for SLS and 4% of TEG. Water was used as solvent.

Skin impedance spectra were measured before and 24 hours after removal of the chemicals. The impedance spectra were collected at 31 logarithmically distributed frequencies from 1 kHz to 1 MHz at five depth settings using an impedance spectrometer provided by SciBase AB, Huddinge, Sweden. The impedance technique and clinical methods are described in detail in [1].

This work was financially supported by SciBase AB, Huddinge, Sweden, and the Karolinska Institute, Stockholm, Sweden

B. Data analysis

There are several approaches in modelling bio-impedance. Traditionally, bio-impedance data have been fitted to equivalent circuits. Several examples of traditional modelling can be found in [2, 3].

Traditional modelling is a theoretical approach with some critical drawbacks. In order to analyse bio-impedance, and emphasise the information from the β-dispersion, in an objective and straightforward way, Ollmar et al. [4], proposed a simple irritation index, IX. A set of four indices (one which is closely related to the old irritation index) were later introduced by Ollmar et al. [5] and used by Nicander et al. [6]. One of the four indices, the magnitude index (MIX) is closely related to the old irritation index. These four indices, based on the magnitude and phase of two frequencies, have been shown to extract a significant amount of the information from full impedance spectra of intact, or nearly intact, skin.

Sethson-Lindholm et al. [7], used multivariate projection methods to classify type 1 diabetics and healthy volunteers using full impedance spectra.

Compared to the traditional, theoretical, way of modelling bio-impedance, the multivariate projection technique is an empirical and statistical approach. The most common multivariate projection tools are principal component analysis (PCA) [8] and partial least squares regression (PLS) [9]. PCA and PLS of bio-impedance is described in detail in [7]. In PCA, sets of orthogonal vectors that describe most of the variance of the data are extracted. The vectors are called scores and loadings. The scores are related to the objects and the loadings are related to the variables. In PLS, the relation between a multivariate data matrix, **X**, set and a set of predictor variables, **Y**, is modelled, **Y** is described by polynomials of **X**. A number of orthogonal PLS components are extracted, in a similar manner as in PCA, that describe most of the variance of **X** and **Y** data and maximise the correlation between **X** and **Y**.

Orthogonal signal correction (OSC) [10] is a data pre-treatment tool that uses multivariate projections. The OSC algorithm finds and removes variations in **X** that are orthogonal to **Y**, i.e. OSC removes unwanted information from **X** that is not correlated to **Y**.

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle Multivariate Regression Model of Impedance of Normal and Chemically Irritated Skin Shows Predictive Ability	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Medical Engineering Karolinska Instituet Huddinge, Sweden	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.

Abstract

Subject Terms

Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU

Number of Pages
4

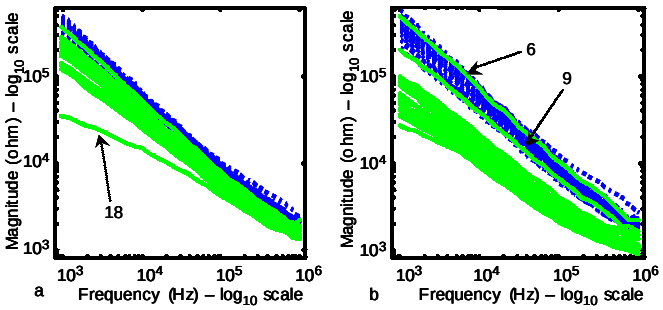

Fig 1. Magnitude, depth setting number 4, measured before (dotted line) and after (solid line) exposure with TEG (a) and SLS (b).

Data handling and visualisation were done using MATLAB 5.0 by MathWorks Inc. Multivariate modelling and OSC filtering were done using SIMCA-P 8.0 by Umetrics AB, Umeå, Sweden.

III. Results

There was a clear irritation effect caused by the chemicals. Raw impedance spectra of the magnitude, depth setting number 4, of normal and irritated skin is visualised in fig. 1. Analysis of variance (ANOVA) of MIX, it was found that there is a 99.9% significant difference between impedance of skin before and after chemical contact. SLS induced more irritation than TEG. MIX, depth setting number four, of the persons before and after chemical exposure are visualised in fig. 2. The straight line is where persons not affected by the chemicals would fall.

There are some readings with deviating impedance, here called outliers; persons number 3, 18, of the TEG site and number 6, 9, and 20 of the SLS site. Some of the outliers are marked in fig. 1 and 2. The cause of the deviating impedance is explained in table I. The outliers were detected using visual inspection of the spectra, inspection of the indices and multivariate projection methods (PCA and PLS) of the full impedance spectra. The outliers are excluded in the following models.

An impedance spectrum is a 2-way data set; every reading is a matrix of frequencies x depth settings. Hence, several readings will produce 3-way data [11], a 3-way array of readings x frequencies x depths. I.e., the structure of the measured data was 21 readings x 31 frequencies x 5 depth settings.

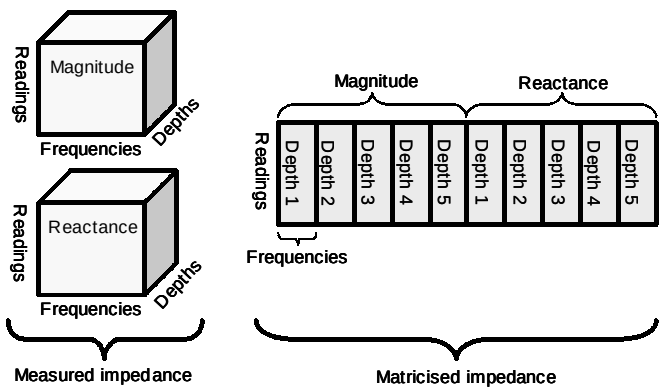

Fig. 3. Structure of the measured 3-way and the matricised impedance.

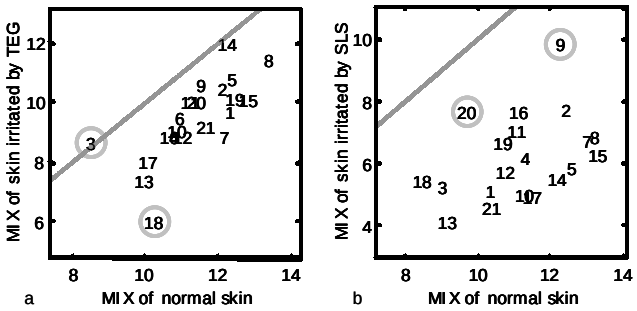

Fig. 2. MIX, depth setting 4, of skin before and after chemically induced irritations with TEG and SLS. Outliers are marked with circles.

Numerically, impedance, Z, is a complex quantity that consists of a real part, R, resistance, and an imaginary part, X, reactance, given by (1). Complex impedance can also be expressed in polar form using the magnitude, |Z|, and phase angle, θ, according to (2).

$$Z = R + iX \quad (1)$$

$$Z = |Z|e^{i\theta}, \quad |Z| = (R^2+X^2)^{0.5}, \quad \theta = \tan^{-1}(X/R) \quad (2)$$

Using common projection techniques (PCA and PLS), it is not possible to model 3-way arrays or complex numbers. The complex impedance was converted to magnitude and reactance and the 3-way structure was matricised, visualised in fig. 3. The final structure of the data was 21 readings x 310 highly correlated variables (31 frequencies x 5 depths x magnitude and reactance).

It was required 4-5 PLS components and about 13 observations to describe the relations between raw impedance and MIX-values of irritated skin. At this stage of analysis, the models were not reliable since the number of observations was not enough to both get a sufficiently large calibration set and to pick an efficiently large test set for validation. In order to reduce unwanted variance and to reduce the number of PLS components, the data was filtered using OSC.

Calibration sets, approximately half of the total number of observations, were selected. The remaining observations were used as external test sets. The matricised impedance of untreated skin was used as **X**-variables and MIX, depth

TABLE I
THE OUTLIERS AND THE CAUSE OF THE DEVIATING IMPEDANCE

Patient no	Site	Cause
3	TEG	The skin did not react when treated with TEG and MIX before the chemical treatment was lower than normal
18	TEG	Low MIX after chemical contact
6	SLS	Experimental mistake
9	SLS	Very high MIX after SLS treatment
20	SLS	The impedance did not fit into the general impedance pattern of the volunteers

setting number one, of the irritated skin impedance was used as **Y**-variable. Two OSC models were calculated; one for the TEG experiment and one for the SLS experiment. The OSC algorithm removed 70% of the variance in the TEG data and 45% of the variance in SLS data.

Analysing the OSC loadings of the two models it was found that the high frequency region contained more unwanted variance, variance orthogonal to MIX, than the low frequency region. This is in agreement with the noise levels in bio-impedance spectra; the noise-level at low frequencies is lower than in the high-frequency region.

The correlation between the skin impedance spectra of skin before chemical exposure and the MIX values of the five depths of the impedance after chemical contact was modelled using PLS. One PLS component for each model was used to describe the relation. The models captured 80-90% of the variance of the **X**-matrix and the **Y**-variables. The cross-validated predictive ability, Q^2, of the models was overall 80-90%. Since the number of readings of the calibration set was limited, Q^2 was calculated using leave-one-out technique [12]. Properties of the two models are listed in table II.

The root mean square error of prediction (RMSEP) is a value of the prediction residuals in original units. It is given by (3), where y_i is the measured and $\hat{y}_i$ is the predicted value of observation i. The predictive ability of the models was good and RMSEP of the external test sets was low (0.74 to 1.09). The depths were highly correlated and RMSEP of the different depth settings (listed in table II) did not differ significantly. Observed vs predicted MIX, depth number four, of the external test sets of the two separate models are visualised in fig. 4a. The straight line is where error-free predictions would fall.

$$\mathrm{RMSEP} = \frac{1}{I}\sum_{i=1}^{I}\sqrt{(y_i - \hat{y}_i)^2} \qquad (3)$$

Variable importance in the projection (VIP) is a value of the contribution of each variable in a model. Variables with large VIP (larger than one) are the most important. VIP of the frequencies are visualised in fig. 5a. Frequencies of the low and middle frequency region (approximately 1 to 100 kHz) are the most important for the regression model. VIP of the SLS model is shifted a bit towards the high frequency region compared to the TEG model. High frequencies (over 400 kHz) seem to be irrelevant for these PLS-models.

TABLE II
PROPERTIES OF THE TWO PLS MODELS

Depth	TEG			SLS		
Setting	R^2	Q^2	RMSEP	R^2	Q^2	RMSEP
1	0.93	0.90	0.74	0.95	0.94	0.82
2	0.87	0.81	0.91	0.93	0.92	0.82
3	0.82	0.77	1.09	0.93	0.91	0.82
4	0.86	0.80	0.84	0.94	0.93	0.88
5	0.84	0.78	0.82	0.93	0.91	0.86

Fig. 5b shows VIP for the five depth settings. There is no significant difference between the depths, the VIP of the depths range from 0.9581 to 1.0041. This indicates that, in this case, the information from depths is less important than the information from the different frequencies.

The VIP of magnitude and reactance is visualised in fig. 4a. The VIP of the reactance is a bit higher than the magnitude. This is an indication that the reactance is a bit more important in predicting the MIX than the magnitude. Further analysis is required to fully understand this phenomenon.

It is evident that all **X**-variables correlate, more or less, and it is possible to reduce the number of measured variables to minimise the size of the data and the sampling time, without affecting the prediction precision of the PLS models.

IV. DISCUSSION

There is a clear relation between bio-impedance of normal and irritated skin. In order to formulate more general conclusions, further experiments with a larger group of test persons is required to validate the method thoroughly. Further experiments in this area will also include additional independent reference methods like transepidermal water loss (TEWL) and visual scoring of healthy volunteers with normal skin and patients with sensitive skin. Additional experiments are in progress.

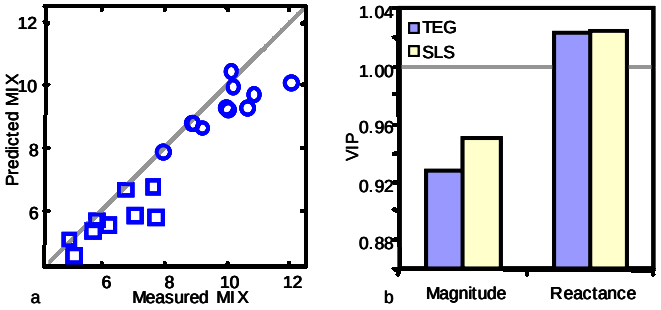

Fig. 4. VIP of magnitude and reactance (a). Measured vs predicted MIX, depth setting number 4, of the TEG (o) and SLS models (□) (b).

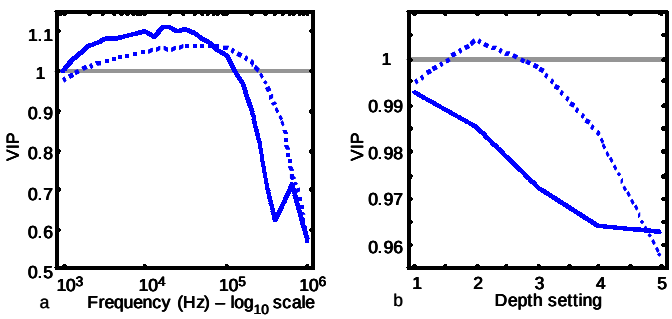

Fig. 5.VIP of the frequencies (a) and the depth settings(b) for the TEG (solid line) and SLS models (dotted line).

Volunteers with abnormal reactions to the chemicals, here called outliers, were detected. This is in agreement with the fact that people with different skin-types or with certain skin diseases react differently to the same exposure of irritative factors. The technology may, in the future, facilitate screening of people before entering professions with high risk of skin problems, such as hairdressers and dentists.

The predictive ability of this model has now been demonstrated for TEG and SLS. Additional studies are required to establish this property for other irritants. One may speculate if this technology also could predict some skin diseases, which might manifest after years, or at least diagnose some diseases at a very early stage, i.e. before any symptoms are discernible.

Ollmar has suggested the term "electronic biopsies" [13], meaning that in the future some diagnostic work might be done not only non-invasively, but also without exposing the patient to test substances.

ACKNOWLEDGMENT

The authors would like to thank the volunteers for making this study possible.

REFERENCES

[1] I. Nicander, "Electrical impedance related to experimentally induced changes of human skin and oral mucosa", *PhD thesis*, Karolinska Institutet, Stockholm, Sweden, 1998.

[2] S. Grimnes and Ø. Martinsen, "Bioimpedance and bioelectricity basics", Academic Press, London, UK, 2000.

[3] K. Foster and H. Schwan, "Dielectric properties of tissues and biological materials: a critical review", *Crit. Rev. Biomed. Eng.*, vol. 17, pp. 25-104, 1989.

[4] S. Ollmar and L. Emtestam, "Electrical impedance applied to non-invasive detection of irritation in skin", *Contact dermatitis*, vol. 27, pp. 37-42, 1992.

[5] S. Ollmar and I. Nicander, "Information in multi frequency measurement on intact skin", *Innovation and technology in biology and medicine*, vol. 16, pp. 745-751, 1996.

[6] I. Nicander, S. Ollmar, A. Eek, B. Lundh Rozell, and L. Emtestam, "Correlation of impedance response patterns to histological findings in irritant skin reactions induced by various surfactants", *British Journal of Dermatology*, vol. 134, pp. 221-228, 1996.

[7] B. Sethson-Lindholm, S. Han, S. Ollmar, I. Nicander, G. Jonsson, F. Lithner, U. Bertheim, and P. Geladi, "Multivariate analysis of skin impedance data in long-term type 1 diabetic patients", *Chemometrics and Intelligent Laboratory Systems*, vol. 44, pp. 381-394, 1998.

[8] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis", *Chemometrics and Intelligent Laboratory systems*, vol. 2, pp. 37-52, 1987

[9] P. Geladi and B. Kowalski, "Partial least-squares regression: a tutorial", *Analytica Chimica Acta*, vol. 185, pp. 1-17, 1986.

[10] S. Wold, H. Antti, F. Lindgren, and J. Öhman, "Orthogonal signal correction of near-infrared spectra", *Chemometrics and Intelligent Laboratory Systems*, vol. 44, pp. 175-185, 1998.

[11] P. Geladi and P. Åberg, "Three-way modelling of a batch organic synthesis monitored by near infrared spectroscopy", *Journal of near infrared spectroscopy*, vol. 9, pp. 1-9, 2001.

[12] S. Wold and L. Eriksson, "Validation tools", in H. van de Waterbeemd ed., "Chemometric methods in Molecular Design", VCH, Weinheim, pp. 309-318, 1995.

[13] S Ollmar, "Making electronic biopsies into a viable future for non-invasive diagnostics with electrical impedance", *Med. Biol. Eng. Comput.*, vol. 37, pp 116-117, 1999.