

## Journal Pre-proofs

Cardio-metabolic risk modeling and assessment through sensor-based measurements

Daniela Giorgi, Luca Bastiani, Maria Aurora Morales, Maria Antonietta Pascali, Sara Colantonio, Giuseppe Coppini

PII: S1386-5056(22)00137-X

DOI: <https://doi.org/10.1016/j.ijmedinf.2022.104823>

Reference: IJB 104823

To appear in: *International Journal of Medical Informatics*

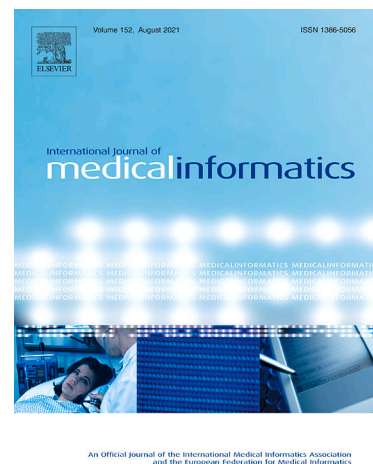
Received Date: 7 February 2022

Accepted Date: 20 June 2022

Please cite this article as: D. Giorgi, L. Bastiani, M.A. Morales, M.A. Pascali, S. Colantonio, G. Coppini, Cardio-metabolic risk modeling and assessment through sensor-based measurements, *International Journal of Medical Informatics* (2022), doi: <https://doi.org/10.1016/j.ijmedinf.2022.104823>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier B.V.



**Cardio-metabolic risk modeling and assessment through sensor-based measurements****HIGHLIGHTS**

- Smart devices give new perspectives for the assessment of cardiometabolic risk in daily life settings.
- Using Structural Equation Models, a three-variable clinical model based on anthropometry, glycolipid factors and vascular function, is defined.
- Sensor measurements about anthropometry, skin auto-fluorescence and endothelial function are taken on subject face using the sensors of the Wise Mirror system.
- By means of Self Organizing Maps, our measures are able to identify subjects at-risk in good agreement with clinical evaluation.

# Cardio-metabolic risk modeling and assessment through sensor-based measurements

Daniela Giorgi<sup>a</sup>, Luca Bastiani<sup>b</sup>, Maria Aurora Morales<sup>b</sup>, Maria Antonietta Pascali<sup>a</sup>, Sara Colantonio<sup>a</sup>, Giuseppe Coppini,<sup>b,\*</sup>

<sup>a</sup>CNR Institute of Information Science and Technologies, Via G. Moruzzi 1, Pisa, 56124, Italy

<sup>b</sup>CNR Institute of Clinical Physiology, Via G. Moruzzi 1, Pisa, 56124, Italy

## Abstract

**Objective:** Cardio-metabolic risk assessment in the general population is of paramount importance to reduce diseases burdened by high morbidity and mortality. The present paper defines a strategy for out-of-hospital cardio-metabolic risk assessment, based on data acquired from contact-less sensors.

**Methods:** We employ Structural Equation Modeling to identify latent clinical variables of cardio-metabolic risk, related to anthropometric, glycolipidic and vascular function factors. Then, we define a set of sensor-based measurements that correlate with the clinical latent variables.

**Results:** Our measurements identify subjects with one or more risk factors in a population of 68 healthy volunteers from the EU-funded SEMEOTICONS project with accuracy 82.4%, sensitivity 82.5%, and specificity 82.1%.

**Conclusions:** Our preliminary results strengthen the role of self-monitoring systems for cardio-metabolic risk prevention.

**Key words:** Cardio-metabolic risk, Risk modeling, Self-monitoring, Smart mirror, Sensor-based measurements, Structural Equation Modeling, Self Organizing Maps

## 1. Introduction

Cardiovascular disease (CVD) represents the world's leading cause of death [1]: the World Health Organization estimates 23.6 million deaths by 2030. Preventing CVD is therefore a main global challenge. In this view, cardio-metabolic (CM) risk refers to those factors that may increase the likelihood of developing vascular events or diabetes. CM risk involves traditional factors included in risk calculators used in clinical practice (e.g., arterial hypertension, dyslipidemia, and smoking) and emerging risk factors (e.g., abdominal obesity, inflammatory profile, and ethnicity) [2]. Noteworthy, most factors can be reduced by improving individual lifestyle.

The identification of at-risk subjects in the general population is of paramount importance to prevent the

development of overt disease and of co-related complications, which bear social and economical consequences [3, 4]. A key issue is to provide people with tools for self-assessing risk factors [5]. Recently, great attention has been paid to eHealth and mHealth applications [6]. Smart devices give new perspectives to CM risk prevention in every-day life settings: prevention is expected to evolve towards smart, individual and proactive strategies particularly focused to lifestyle improvement.

In this paper, we define a strategy for CM risk assessment for primary prevention in the general population. Our strategy leverages on statistical modeling, data analysis and advanced sensor-based monitoring technology, and can be implemented as part of a non-invasive monitoring system placed at home or other daily-life settings, such as gyms and chemist's shops. A reliable at-home monitoring system for CM risk would reduce the number of people in care offices (decreasing the burden on medical professionals), and increase adherence with individually-tailored prevention actions.

Our approach consists of two pathways. First, we define a clinical model of CM risk factors, based on up-

\*Corresponding author

Email addresses: daniela.giorgi@isti.cnr.it (Daniela Giorgi), luca.bastiani@ifc.cnr.it (Luca Bastiani), maria.aurora.morales@ifc.cnr.it (Maria Aurora Morales), maria.antonietta.pascali@isti.cnr.it (Maria Antonietta Pascali), sara.colantonio@isti.cnr.it (Sara Colantonio), giuseppe.coppini@ifc.cnr.it (Giuseppe Coppini)

Preprint submitted to International Journal of Medical Informatics

List of acronyms used in the paper.

CVD	Cardiovascular disease
CM	Cardiometabolic
ML	Machine Learning
SEM	Structural Equation Modeling
SOM	Self Organizing Maps
DT	Decision Trees
RF	Random Forests
$k$ -NN	$k$ -Nearest Neighbour
RHI	Reactive Hyperemia Index
BMI	Body Mass Index
LDL	Low-Density Lipoprotein
HbA1c	Hemoglobin A1c
AGE	Advanced Glycation End-products
fRBC	fraction of Red Blood Cell Count
UV	Ultraviolet
LED	Light Emitting Diode
MSI	Multispectral Imaging
SE	Standard Error
RMSEA	Root Mean Square Error of Approximation
SRMR	Standardized Root Mean Residual
CFI	Comparative Fit Index
TLI	Tucker-Lewis Index

to-date clinical knowledge and standard clinical practice. Then, we define a set of measurements closely related with clinical risk factors, which can be evaluated at home through non-contact sensors. We demonstrate that our sensor-based measurements can recognize at-risk subjects, and provide a proof-of-concept for a personalized strategy for risk prevention.

To define the clinical model of CM risk, we collect clinical data on a population of 68 healthy subjects, and carry out Confirmatory Factor Analysis via Structural Equation Modeling (SEM) [7]. The analysis confirms the presence of three latent variables corresponding to different risk categories, namely, risk related to anthropometric factors, glycolipid function, and vascular function. SEM is gaining momentum in disciplines such as psychology, social and economic sciences, and also medicine [8], as a technique to analyse conceptual models and quantify the relationships among a network

of factors. As opposed to black-box machine learning techniques, SEM explains how single factors contribute to intermediate latent variables and to the final risk outcome.

After defining the clinical model, we select a set of sensor-based measurements which are closely related to the latent variables of the clinical model, and which can be evaluated non-invasively in the context of self-monitoring at home. The measurements are taken on facial features, according to a *semeiotic* model of CM risk [9]. We show that the sensor-based measurements have significant correlation with the latent variables from clinical parameters. Therefore, they can be used in place of clinical parameters for non-invasive self-monitoring at home. Furthermore, we use statistical analysis and Self Organizing Maps (SOMs) to show that our measurements are able to identify subjects at-risk, thus supporting the development of self-monitoring systems that warn individuals about the onset of CM risk, enable them to act on individual risk factors, and trigger medical examination when needed.

Remarkably, our CM risk monitoring strategy is explainable by design: as our sensor-based measurements correlate with latent clinical variables identified via SEM, they inherit the interpretability of the underlying clinical model.

To sum up, our main contributions are:

- defining a clinical model of CM risk. While there are many studies on CM risk factors, our study of associations via SEM analysis can contribute to shed light on the multifactorial etiology of CM risk;
- defining sensor-based measurements that correlate with clinical parameters and that can be non-invasively acquired at home or other out-of-hospital settings;
- demonstrating that sensor-based measurements are able to identify at-risk subjects, in good agreement with clinical evaluation;
- a proof-of-concept about the potential of integrating a multi-sensing platform with proper data modeling strategies, for the definition of CM risk indicators in the context of personalized monitoring and primary prevention in the general population.

The paper layout is as follows. Section 2 discusses the state of the art about CM risk assessment. Section 3 introduces the dataset. Section 4 describes the SEM model based on standard clinical data and the sensor-based measurements. Section 5 provides results about the SEM model estimation and its consistency with clinical evaluation of CM risk, the correlation of

109 sensor-based measurements with SEM latent variables, 157  
 110 the clinical evaluation of CM risk, and the recognition 158  
 111 of at-risk conditions by SOMs. Conclusions are drawn 159  
 112 in Section 6. 160

## 113 2. State of the art

### 114 2.1. Cardio-metabolic risk indicators

115 Several validated risk charts are reported in the med- 161  
 116 ical literature [14–20]. Most risk scores use standard 162  
 117 CVD risk factors (age, sex, smoking, blood pressure and 163  
 118 cholesterol); some also incorporate advanced markers 164  
 119 on metabolic or homeostasis processes. As opposed to 165  
 120 existing risk scores that tend to capture specific features, 166  
 121 the risk model in the present paper is multi-faceted, as 167  
 122 it takes into account the whole spectrum of CM risk, 168  
 123 including both CVD and type2-diabetes. 169

124 A recent survey [21] debates the use of CM risk 170  
 125 scores in clinical practice on the basis of clinical out- 171  
 126 comes. While the use of risk charts in sporadic visits 172  
 127 by specialized medical professional may have a lim- 173  
 128 ited positive effect, we hypothesize that a continuous 174  
 129 and personalized assessment may guarantee a thorough 175  
 130 monitoring of risk factors and a timely delivery of alerts. 176  
 131 Several solutions have been devised so far for the re- 177  
 132 mote monitoring of chronic patients [10, 22, 23], while 178  
 133 few attempts target CM risk prevention in healthy sub- 179  
 134 jects [11]. None of these works has yet defined a per- 180  
 135 sonalized risk assessment tool. 181

136 In this paper, we present a proof-of-concept for 182  
 137 a personalized preventative solution based on self- 183  
 138 monitoring, through measurements computed at home 184  
 139 via contact-less, non-invasive sensor measurements. 185  
 140 Table 1 compares our proposal with the works dis- 186  
 141 cussed in this section. 187

### 142 2.2. Machine learning for cardio-metabolic risk assess- 188 143 ment 189

144 Recent works have tried to improve the accuracy of 190  
 145 existing CM risk scores via Machine Learning (ML). 191  
 146 The authors of [12] frame risk prediction as a classi- 192  
 147 fication problem and compare three ML methods with 193  
 148 the HellenicSCORE, on a dataset that comprises demo- 194  
 149 graphic, metabolic and biometric variables. The ML 195  
 150 methods are  $k$ -Nearest Neighbours ( $k$ -NNs), Decision 196  
 151 Trees (DTs) and Random Forests (RFs). ML meth- 197  
 152 ods do not outperform the HellenicSCORE. The authors 198  
 153 also comment on  $k$ -NNs and Random Forests classifiers 199  
 154 being not easily intelligible, and making it hard to ex- 200  
 155 plain classification results. On the other hand, Deci- 201  
 156 sion Trees are easier to understand, yet more simplistic 202  
 203  
 204

157 when compared to the other models. Another study [13] 205  
 158 investigates whether ML can improve the accuracy of 206  
 159 risk prediction within a large general primary-care pop- 207  
 160 ulation. The authors compare the prediction accuracy 208  
 161 of the ACC/AHA index [24] against logistic regression, 209  
 162 Random Forests, gradient boosting machines, and artifi- 210  
 163 cial neural networks (ANN). The results show ML algo- 211  
 164 rithms outperform the ACC/AHA index. Nonetheless, 212  
 165 the best performance is obtained by an artificial neural 213  
 166 network, which suffers from the so-called “black-box” 214  
 167 effect, despite the use of explanatory visualization tech- 215  
 168 niques. 216

169 On the contrary, we model CM risk to be explain- 217  
 170 able by design, thanks to the use of SEM, a data-driven 218  
 171 approach suitable to identify latent variables and their 219  
 172 influence in an easily interpretable way. 220

### 173 2.3. SEM techniques

174 SEM is a technique to discover pathways of associ- 221  
 175 ations between latent and observed variables, by tak- 222  
 176 ing into account collinearities in the data. We refer the 223  
 177 reader to [7] for a comprehensive description. 224

178 Khodarahmi et al [25] use SEM to assess the as- 225  
 179 sociation of adherence to a healthy-eating index with 226  
 180 socio-demographic factors, psychological characteris- 227  
 181 tics, and CM risk factors among obese individuals. 228  
 182 Lewlyn et al. [26] reveal via SEM the positive asso- 229  
 183 ciation of cigarettes smoked per day, alcohol consumed 230  
 184 per week, and diastolic blood pressure with hyperten- 231  
 185 sion and coronary heart disease. Shakibaei et al. [27] 232  
 186 use SEM to investigate the integration of standard med- 233  
 187 ical data to assess CM risk in clinical settings. 234

188 SEM techniques have been proven effective and ro- 235  
 189 bust on datasets of relatively small size. Another major 236  
 190 advantage of SEM is that it can be used when no su- 237  
 191 pervisory information is available on the data, as in our 238  
 192 context. 239

## 193 3. Dataset

194 To set up the clinical model of CM risk, and then 240  
 195 test the sensor-based self-monitoring strategy, we col- 241  
 196 lected data about a population of 75 volunteer subjects 242  
 197 in overall healthy conditions. Being healthy does not 243  
 198 exclude the presence of potential CV risk factors, and 244  
 199 we aim indeed at primary prevention in the general pop- 245  
 200 ulation. In particular, our population was chosen on the 246  
 201 basis of lack of physical and mental disease at least for 247  
 202 6 months before enrollment; care was taken to exclude 248  
 203 subjects under any sort of medical treatment or previ- 249  
 204 ous autoimmune or neoplastic disease and specifically 250

Table 1: Positioning our proposal within the state of the art. Top: with respect to delivered output and target population. Bottom: with respect to methodological aspects and reference clinical indicators.

Reference	Delivered output	Target
[10]	Decision Support System	Chronic patients, Follow up
[11]	Digital platform	Chronic patients, Secondary prevention
Ours	Multisensory platform, CM Risk model	Healthy subjects, Primary prevention

Reference	Methods	Reference clinical indicators	Target
[12]	$k$ -NN, DT, RF	HellenicSCORE	Primary prevention
[13]	Logistic Regression, RF, ANN	ACC/AHA index	Primary prevention
Ours	SEM, SOM	Novel, multi-faceted clinical model of CM risk	Primary prevention

Table 2: Clinical parameters for CM risk assessment and their statistics, grouped by risk factors.

			Male (30)	Female (38)	Total	P-value
<b>Anthropometric factors</b>		<25	23.3%	36.8%	30.9%	
	Body Mass Index Class	25 to 29.9	50.0%	31.6%	39.7%	0.278
		>30	26.7%	31.6%	29.4%	
		Waist Circumference (cm)	25 to 29.9	50.0%	31.6%	39.7%
	Hip Circumference (cm)	25 to 29.9	50.0%	31.6%	39.7%	0.787
	Fat Mass	25 to 29.9	50.0%	31.6%	39.7%	0.001
<b>Glycolipid factors</b>	Cholesterol Levels	mean $\pm$ sd	195.8 $\pm$ 33.1	200.2 $\pm$ 42.6	198.3 $\pm$ 38.5	0.645
	LDL (mg/dl)	mean $\pm$ sd	122.1 $\pm$ 28.7	122.3 $\pm$ 32.9	122.2 $\pm$ 30.9	0.974
	Glucose (mg/dl)	mean $\pm$ sd	95.3 $\pm$ 12.4	88.5 $\pm$ 9.8	91.5 $\pm$ 11.5	0.140
	HbA1c (mmol/mol)	mean $\pm$ sd	36.6 $\pm$ 3.5	36.4 $\pm$ 4.0	36.5 $\pm$ 3.8	0.875
	Triglycerides	mean $\pm$ sd	117.8 $\pm$ 61.3	92.6 $\pm$ 47.9	103.7 $\pm$ 55.2	0.061
<b>Vascular function</b>	Reactive Hyperemia Index	mean $\pm$ sd	2.1 $\pm$ 0.5	2.4 $\pm$ 0.7	2.3 $\pm$ 0.6	0.118

those with known systemic hypertension, hypercholesterolemia, diabetes. Subjects with increased body mass index alone were not excluded keeping in mind that the category of *healthy obese* exist. In 68 enrolled subjects full data were available for the present study. At baseline, all subjects underwent a complete medical history. Then, a physical examination followed. The data analysed in this paper were collected once for each subject. The characteristics of the study population and the data collected are described in Appendix A.

Our working hypothesis is that CM risk can be described in terms of three main risk factors: anthropometric factors, glycolipid factors, and vascular function. Table 2 reports the set of clinical parameters used in the present study for CM risk assessment and some basic

statistical values, with the parameters grouped according to the three risk factors above mentioned:

- Anthropometric factors: four anthropometric parameters that are sensible to obesity and overweight, and that are commonly used in clinical practice;
- Glycolipidic factors: abnormal lipid metabolism and hyper-glycaemia, which are recognized CM risk factors;
- Vascular function: the Reactive Hyperemia Index (RHI) measured by pulse amplitude tonometry [28], to measure endothelial dysfunction, which is a major physio-pathological mechanism correlated with CM risk factors, leads towards coronary artery disease, and is involved in several dis-



ease processes (e.g. hypertension, hypercholesterolemia, and diabetes).

## 4. Methods

Section 4.1 describes our modeling strategy to define latent variables on top of the clinical measurements listed in Table 2. Then, Section 4.2 defines the sensor-based counterparts for the clinical latent variables, to be computed in the context of CM risk self-monitoring.

### 4.1. Identifying latent variables in CM risk

In terms of SEM, a model of CM risk can be based on a set of linear equations that relate observed clinical parameters to latent variables representing different risk components [7].

As observed above, we focus on CM risk related to anthropometric factors, glycolipid factors, and vascular function. Since for vascular function we have a single observed parameter (RHI), no latent variable is introduced. For the other two risk components (anthropometric and glycolipidic), we model the relations between observations and latent clinical variables as in Figure 1, by defining the two latent variables:

- Anthropometric factors variable, depending on Body Mass Index (BMI), Waist Circumference, Hip Circumference, and Fat Mass;
- Glycolipidic factors variable, depending on LDL Cholesterol level, Glucose, Glycated hemoglobin, and Triglycerides.

Using the notation in Table 3, we denote with  $\{\lambda_i\}_{i=1}^4$  and  $\{\gamma_i\}_{i=1}^5$  the clinical parameters, with  $\Lambda$  and  $\Gamma$  the corresponding latent variables. Therefore, we can write the following structural equations:

$$\lambda_i = b_i + c_i\Lambda + \epsilon_i, \quad i = 1, \dots, 4$$

$$\gamma_j = d_j + e_j\Gamma + \epsilon_j, \quad j = 1, \dots, 5$$

where  $b_i, c_i, d_j, e_j$  are the model coefficients and  $\epsilon$ s are noise terms.

In Section 5 we estimate the model coefficients by fitting the model to the population described above, and analyse how the single observed variables (corresponding to clinical parameters) contribute to each latent risk factor.

### 4.2. Non-invasive self-monitoring of CM risk via sensor-based measurements

We propose a set of sensor-based measurements as counterparts for the two latent clinical variables (anthropometric and glycolipidic factors variables) and for the

Table 3: Symbol convention for SEM modelling

Variable name	Latent Variable	Measurement
<b>Anthropometric factors</b>		$\Lambda$
BMI Class		$\lambda_1$
Waist Circ.		$\lambda_2$
Hip Circ.		$\lambda_3$
Fat Mass		$\lambda_4$
<b>Glycolipidic factors</b>		$\Gamma$
Cholesterol Levels LDL		$\gamma_2$
Glucose		$\gamma_3$
HbA1c		$\gamma_4$
Triglycerides		$\gamma_5$

vascular function parameter (RHI). The measurements can be non-invasively evaluated via a self-monitoring device. We use the multi-sensing system developed in the context of the European project SEMEOTICONS. The system is called *Wize Mirror*, as it has the appearance of a mirror to easily fit into daily-life settings (Figure 2). The *Wize Mirror* includes a 3D acquisition module with a low cost depth sensor for face detection, reconstruction, and morphometric analysis; and a multispectral imaging (MSI) module with five compact monochrome cameras with band-pass filters at selected wavelengths and two computer-controlled LED light sources [29]. Three prototypes of the *Wize Mirror* were deployed in three clinical sites (Pisa, Milan and Lyon), where sensory data were acquired.

Our sensor-based measurements for CM risk assessment derive from the face semeiotic model of CM risk in [9]. They are listed below, grouped according to the three risk categories identified in the previous sections. All measurements are non-invasive and contact-less:

- Anthropometric measurement (*Wize Mirror MorphoE*): We compute *Wize Mirror MorphoE* as the maximal length of curves resulting from intersecting the 3D face surface and a set of spheres centered in the nose tip and with increasing radius (Figure 3.a). This measure has been shown to correlate with standard weight-related measurements (weight, body mass index, waist and neck circumference), and therefore is an indicator of overweight and obesity [30];
- Glycolipidic measurement (*Wize Mirror AGE*): *Wize Mirror AGE* quantifies AGE (Advanced Glycation End-products) deposits of skin tissue, which are favoured by metabolic alterations due to diabetes [31]. AGE in sub-cutaneous layer can be detected via autofluorescence stimulated by UV light [32]. We use the technique in [33], based on the ac-

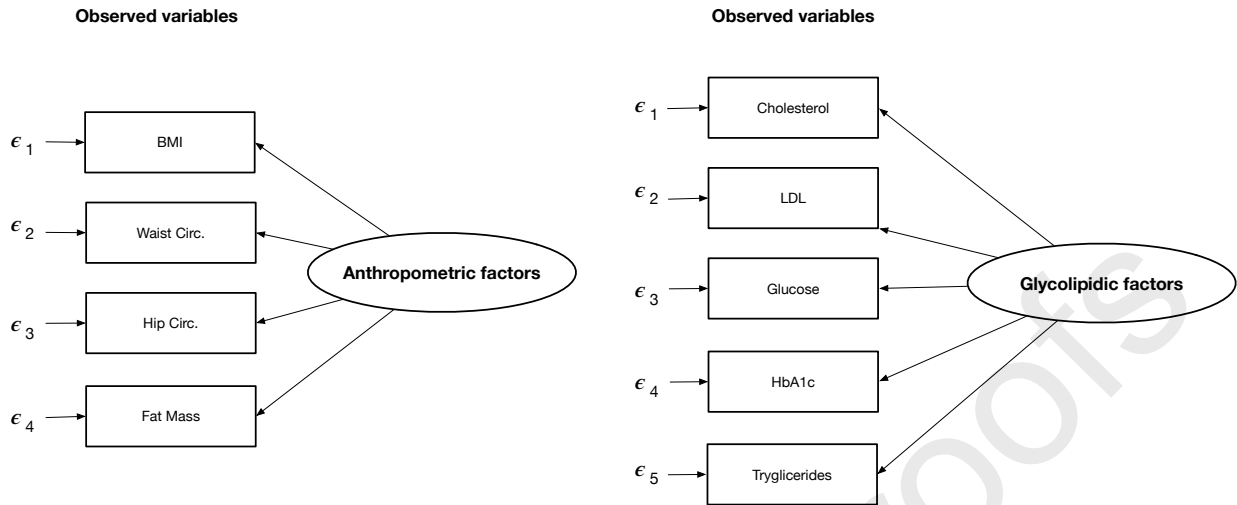


Figure 1: Latent clinical variables (ellipses) and their relation to clinical observations (rectangles).  $\epsilon_{i,ss}$  are noise terms.



Figure 2: The Wize Mirror prototype.

quisition of multispectral images of the face during UV exposure (Figure 3.b);

- Endothelial Dysfunction measurement (*Wize Mirror ENDO*): Our sensor-based measurement *Wize Mirror ENDO* is based on the analysis of microcirculatory blood flow after local heating [34]. Changes in skin fraction of Red Blood Cell Count (fRBC) during local heating are related to reactive hyperemia and can be used as indicators of endothelial function [35]. fRBC can be measured reliably using MSI, after heating the face skin to the temperature of 39°C for about 10 minutes through a computer-controlled heater and an IR thermometer measuring skin temperature.

In Section 5.2 we demonstrate our measurements are positively correlated with the clinical variables, and that

they can identify at-risk subjects in our population. The ability to discriminate between normal and risk conditions is assessed via Self-Organizing Maps (SOMs) [36]. Details on SOMs are given in Appendix B.

## 5. Results

### 5.1. SEM model

#### Estimation of SEM coefficients

We estimate the model coefficients on the SEMEOTI-CONS' data-set. The values of the standardized regression coefficients are reported in Table 4. The regression coefficients show how the single observed variables (corresponding to clinical parameters) contribute to each latent risk factor. The most important predictors for the anthropometric factor score were Body Mass Index (standardized regression coefficient  $\beta = 0.939$ , standard error  $SE = 0.027$ , significance  $p < 0.0001$ ) and Hip Circumference ( $\beta = 0.829$ ,  $SE = 0.043$ ,  $p < 0.0001$ ). For the glycolipid factor score the HbA1c (mmol/mol) ( $\beta = 0.741$ ,  $SE = 0.093$ ,  $p < 0.0001$ ), and the Cholesterol levels ( $\beta = 0.655$ ,  $SE = 0.098$ ,  $p < 0.0001$ ) were the most relevant predictors. The Structural Model Fit indices (Root Mean Square Error of Approximation, RMSEA; Standardized Root Mean Residual, SRMR; Comparative Fit index, CFI; Tucker-Lewis Index, TLI) indicate that the proposed models fit the data adequately [37].

#### Model evaluation

To check the consistency of SEM-derived factor scores with clinical findings, we test if the factor scores



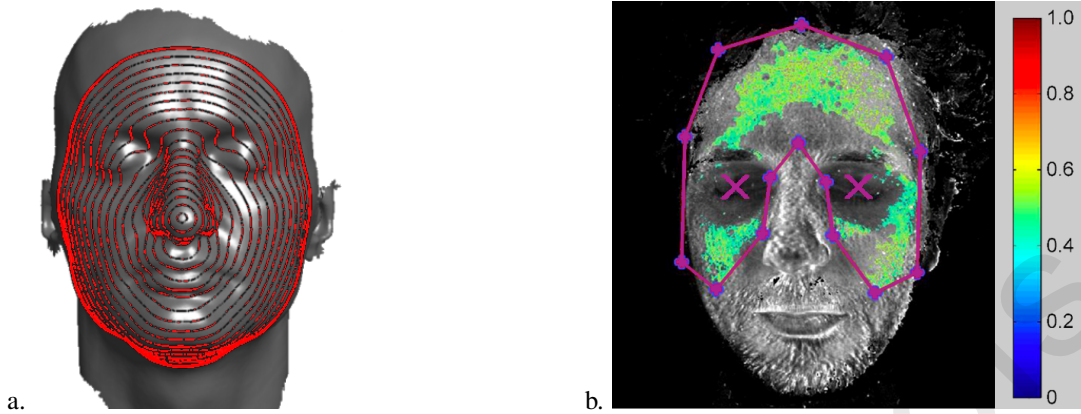


Figure 3: a. The set of curves used to compute the Wise Mirror MorphoE, indicative of fatness or obesity. b. Auto-fluorescence map obtained by UV light exposure. The skin in forehead and cheeks is particularly responsive to the stimulation due to the accumulation of AGEs.

Table 4: Regression coefficient of the SEM model built on top of clinical parameters.

Factor	Observed variables	Regression coefficients			Standardized Regression Coeff. (95% Confidence Interval)	Fit Indices			
		Standardized Regression Coeff.	SE	P >  z		RMSEA	SRMR	CFI	TLI
Anthropometric	Body Mass Index (BMI)	0.939	0.027	0.000	0.885 to 0.993	0.075	0.004	0.935	0.917
	Waist Circ.	0.787	0.054	0.000	0.681 to 0.893				
	Hip Circ.	0.829	0.043	0.000	0.745 to 0.914				
	Fat Mass (Bod Pod)	0.684	0.071	0.000	0.545 to 0.825				
Glycolipid	Cholesterol Levels	0.655	0.098	0.000	0.462 to 0.847	0.098	0.042	0.985	0.963
	LDL (mg/dl)	0.643	0.099	0.000	0.448 to 0.838				
	Glucose (mg/dl)	0.650	0.104	0.000	0.447 to 0.854				
	HbA1c (mmol/mol)	0.741	0.093	0.000	0.558 to 0.924				
	Triglycerides	0.565	0.113	0.000	0.344 to 0.786				

are able to place the subjects in our population in different CM risk categories. A cardiologist grouped the 68 subjects into three classes of CM risk: no risk (green), mild to moderate risk (yellow), high risk (red). The grouping was performed for each of the three latent variables: anthropometric, glycolipidic, and vascular function. The colorization is based on the clinical parameters related with risk factor: green if all parameters fall within normal limits; yellow if at least one parameter is slightly outside the upper limit; and red if at least one parameter is well above upper limits. It is worth noting that Vascular function has two groups (red and green) only. This is due to the fact that a single threshold is used in clinical practice for RHI. Fourteen subjects were classified as green, fourteen as yellow, and forty as red.

The box and whiskers plots in Figure 4.a show how the latent variables identified via SEM were able to discriminate subjects in different risk categories. This confirmed the ability of SEM to correctly identify latent

variables.

## 5.2. Evaluation of sensor-based measurements

### Correlation with clinical risk factors

We analyse the correlation between sensor-based measurements and the latent variables from clinical parameters. All the three measurements have significant positive correlation with their clinical counterparts, with  $p$ -value less than  $10^{-2}$ . In particular, the anthropometric sensor measurement Wise Mirror MorphoE has a Pearson correlation of 0.559 with the anthropometric clinical factor score; the glycolipidic sensor measurement Wise Mirror AGE has a Pearson correlation of 0.349 with the glycolipid clinical factor score; the endothelial dysfunction sensor measurement Wise Mirror ENDO has a Pearson correlation of 0.648 with the endothelial dysfunction clinical factor score. The correlation coefficients and their statistical significance for all three sensor-based measurements show their suit-

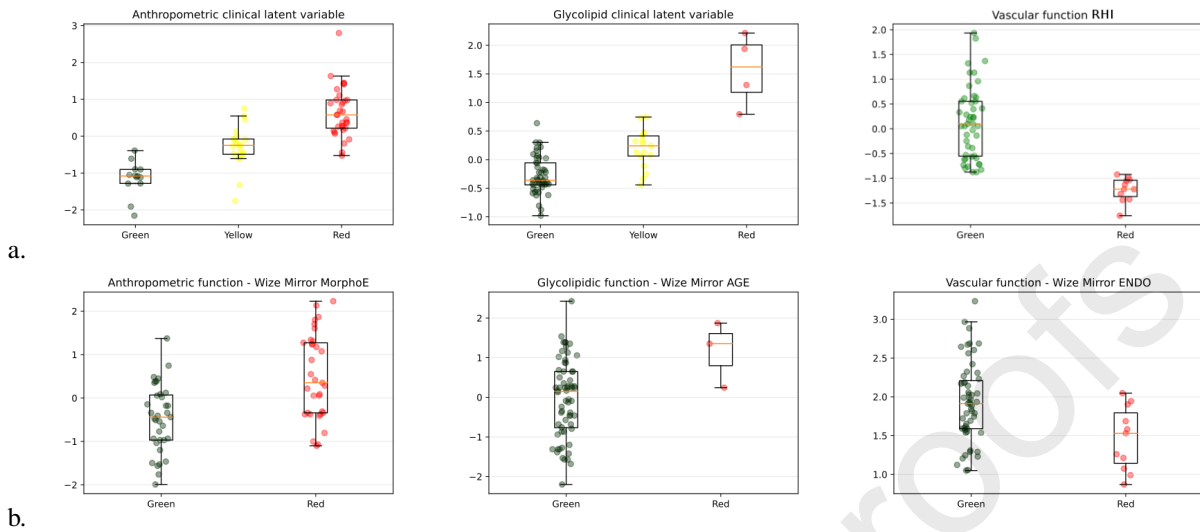


Figure 4: a.) Box and whiskers plots for clinical variables (two latent variables and RHI) on the study population. The clinical variables are able to separate subjects having no risk (green) from subjects with moderate (yellow) or high (red) risk. It worth noting that Vascular function has two groups (red and green) only, because a single threshold is commonly accepted in clinical practice for RHI. The reader should notice that RHI is decreasing for impaired endothelial function. b.) Box-plots of sensor-based descriptors (Wize Mirror MorphoE, Wize Mirror AGE, Wize Mirror ENDO) for risk (red) and no-risk groups (dark green). The dark green group includes both green and yellow group for clinical data

393 ability in place of clinical parameters for non-invasive  
394 self-monitoring.

#### 395 *Usefulness for CM risk self-monitoring*

396 We investigate whether sensors-based measurements  
397 are able to assess individual CM risk, and thus trigger  
398 proper warnings. First, we check whether the sensor  
399 measurements are able to identify people at risk. As  
400 we did with clinical factor scores, we analyse box and  
401 whiskers plots on the study population, for each sensor  
402 measurement. In order to provide a clear cut separation  
403 between subjects with and without one or more CM risk  
404 factors, the sensor-based measurements are categorized  
405 in two groups only: dark green (including people with  
406 normal or slightly outside normal values, i.e., includ-  
407 ing green and yellow subjects in the previous classifi-  
408 cation) and red (absolutely outside normal range, same  
409 as in the previous classification). The box and whiskers  
410 plots of sensor-based measurements are shown in Fig-  
411 ure 4.b. Median values are significantly different in the  
412 two groups for all three measurements. Though, some  
413 overlapping exists between the groups for MorphoE and  
414 ENDO. This was expected, as our population mainly in-  
415 cludes healthy subjects, and those with history or cur-  
416 rent overt cardiovascular diseases and diabetes were ex-  
417 cluded.

#### 418 *SOM analysis*

419 We train 2D Self Organizing Maps on the sensor-  
420 based measurements Wize Mirror MorphoE, Wize Mir-  
421 ror AGE and Wize Mirror ENDO. The training is run  
422 ten times with random weight initialization. We refer  
423 to a SOM with  $7 \times 7$  units, which is coherent with the  
424 data set size and exhibits a good compromise between  
425 data representation and overall accuracy in recognizing  
426 different risk condition.

427 Figure 5.a depicts the distribution of each weight di-  
428 mension in the network space (weight-plane maps). A  
429 clear spatial arrangement of weight values has emerged  
430 after training.

431 To assess the discrimination capabilities of the net-  
432 work in discriminating different risk conditions, we ex-  
433 amine the distribution of winning units with respect to  
434 different data categories. As before, we consider dark  
435 green and red subjects, and evaluate the hit maps for  
436 the two group. In the left panel of Figure 5.b we have  
437 the hits map for normal (dark green) subjects, while in  
438 the middle panel the hit map for at risk (red) subjects  
439 is shown. The two groups of responder units are rather  
440 separated and suggest that the network can discriminate  
441 between the two different risk conditions.

442 Using a majority voting scheme [38], we labeled the  
443 units as representative of the dark green group (lower  
444 CM risk) and red group (higher risk). The resulting la-  
445 belling is reported in the right panel of Figure 5.b. Ac-

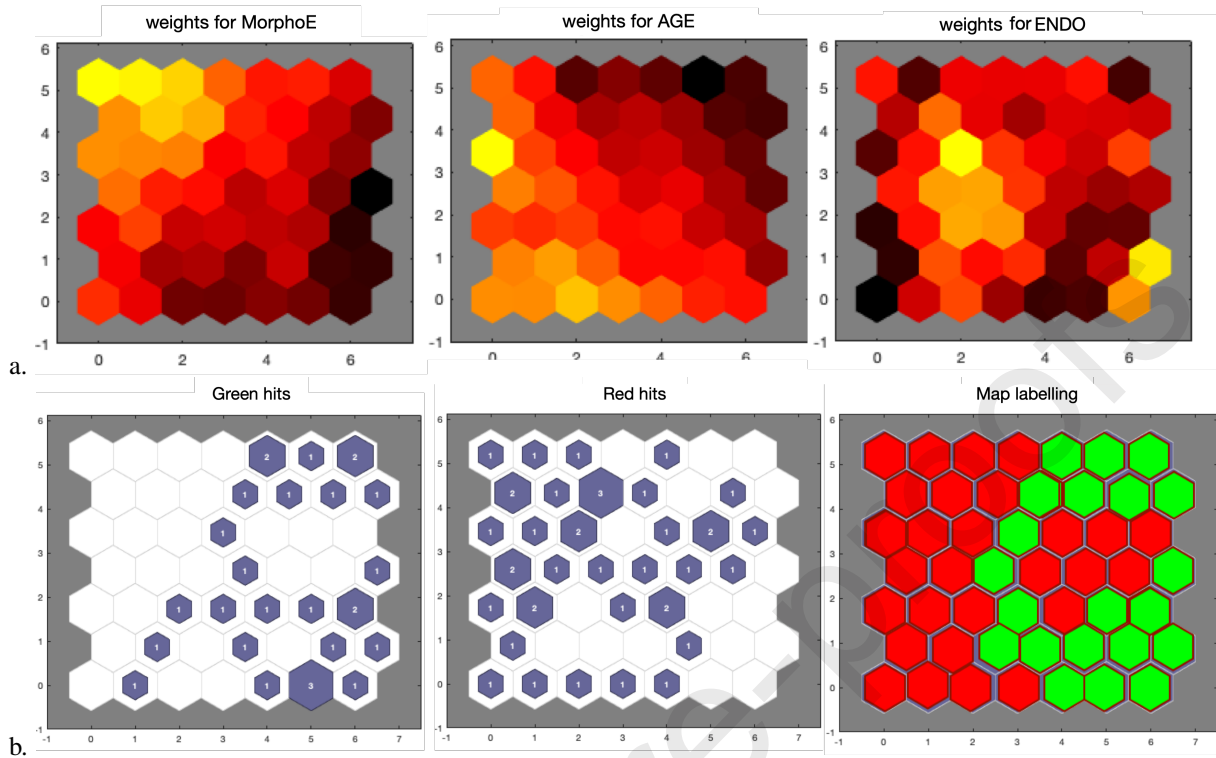


Figure 5: a) Maps of SOM weight components. Darkest colors indicate smallest values while light colors denote largest ones. b) Left: hits map for green subjects, middle: hit map for Red subjects, right: map labeling as obtained using majority voting.

446 cording to this labelling scheme, we observe the classifica-  
 447 tion performance detailed in Table 5 below. Accuracy,  
 448 sensitivity and specificity are all larger than 82%, denot-  
 449 ing good recognition capabilities of subjects at risk.

450 To evaluate SOM behaviour with respect to size, Ta-  
 451 ble 6 reports the performance for SOMs of varying di-  
 452 mensions, coherent with the dataset size: from  $5 \times 5$  to  
 453  $9 \times 9$  units. While maps with size below  $6 \times 6$  are less  
 454 accurate, maps with dimension  $7 \times 7$  or higher show  
 455 better performance. The  $7 \times 7$  map has a number of  
 456 units which guarantees good accuracy while containing  
 457 the risk of overfitting, given the number of subjects in  
 458 our dataset.

## 459 6. Discussions and Conclusions

460 ICT technologies can support efficient strategies  
 461 against the spread of CVD and CM risk, by integrating  
 462 multi-sensing platforms and data modelling to derive a  
 463 personalized evaluation of risk conditions. In this work,  
 464 we use data from the EU project SEMEOTICONS to  
 465 provide a proof of concept of a sensor-based strategy to  
 466 recognize the presence of one or more CM risk factors  
 467 in individuals. The aim is to increase the awareness in

Table 5: Confusion matrix for risk classification by SOM.

Clinical evaluation	Lower risk	Higher risk	Total
SOM classification			
Lower risk	23	7	30
Higher risk	5	33	38
Total	28	40	68

True Positives:	33
False positives:	5
True negatives:	23
False negatives:	7
Accuracy:	0.824 (23+33)/68
Sensitivity	0.825 33/40
Specificity	0.821 23/28

468 apparently healthy subjects about risk factors for dis-  
 469 eases with high rates of morbidity and mortality. Our  
 470 approach is based on two different pathways. From a

Table 6: Results for risk classification by SOM, according to different map sizes.

Size	Accuracy	Sensitivity	Specificity
5 × 5	0,676	0,700	0,643
6 × 6	0,779	0,775	0,786
7 × 7	0,824	0,825	0,821
8 × 8	0,838	0,850	0,821
9 × 9	0,853	0,850	0,857

clinical view point, data collected from volunteers are used to set a simple model of the main risk factors, represented by clinical latent variables. This model is implemented according to the SEM methodology, which is explainable by design, and works well with limited data and no supervisory information. The model consistency with clinical findings is qualitatively reported. From the individual monitoring view point, we adopt sensor-based measurements of face signs. We demonstrate that they are closely related to the latent variables of the clinical model, and that they can identify high-risk subjects with respect to both single risk factors and overall risk.

Our results are preliminary, due to the limited number of sensor-based measurements tested, the use of a prototype for their acquisition, and the relatively small size of the dataset. Nevertheless, our results are promising, especially in light of the fact that identifying at-risk subjects among individuals in overall healthy conditions is not an easy task.

In the future, we plan to include additional sensor-based measurements, related to both physical and psychological aspects relevant to CM risk (e.g., blood pressure data, pulse-oxymetry, heart rate and heart rate variability, facial signs of stress and anxiety). This is expected to improve the sensitivity and specificity of our risk assessment procedure [39, 40]. Another line of research is the stability of the considered analysis to sensor measurement faults [41]. Moreover, encouraged by the results on the dataset from the SEMEOTICONS project, we plan to enlarge the population size. This would also allow us to experiment with different data analysis and learning techniques.

Our results show that a sensor-based, non-invasive assessment of CM risk is feasible for primary prevention. Therefore, our findings contribute to strengthen the role of technology and data modeling for out-of-hospital individual monitoring. In the era of precision medicine, the approach we presented may provide “the right treat-

ment to the right person at the right time”.

## 7. Summary table

- CVD represents the world’s leading cause of death. CM risk refers to those factors that may increase the likelihood of developing CVD or diabetes.
- Smart devices give new perspectives for the assessment of CM risk in every-day life settings.
- A clinical model of CM risk is derived using the data from the population of the EU project SEMEOTICONS (68 healthy subjects). Using SEM method three variable are defined relating to Antropometric factors, Glycolipid factors and Vascular function. In the same population, sensor measurements related to antropometry, skin auto-fluorescence and endothelial function are taken on subject face using the sensors of the Wize Mirror system [29].
- The sensor-based measurements have significant positive correlation with the latent variables from clinical parameters. Using SOMs, our measures are able to identify subjects at-risk in good agreement with clinical evaluation.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Acknowledgment

The authors thank prof. Martine Laville and Prof. Julie-Anne Nazare from the Human Nutrition Research Centre Rhône-Alpes (CRNH, Lyon, France), Renata De Maria MD from the Milan branch of the CNR Institute of Clinical Physiology, and all the SEMEOTICONS project partners for their contribution to data collection.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Dataset

The data come from an acquisition campaign in the context of the EU-funded project SEMEOTI-CONS (FP7 Project - SEMEiotic Oriented Technology for Individual's CardiOmetabolic risk self-assessment and Selfmonitoring, <http://www.semeoticons.eu>, GA. 611516). Volunteers were recruited through local advertisements or during an outpatient visit in one of three clinical centers (Pisa, Lyon and Milan). Subjects were considered eligible according to the following inclusion criteria: age in-between 25 and 60 years; willing to participate in the study; overall healthy conditions, and under no medical treatment at study inclusion. As we target primary prevention in the general population, study exclusion criteria were represented by history or current overt cardiovascular or cerebrovascular disease and/or diabetes. Nevertheless, CM risk factors were not assessed before enrollment. Ethical approval for the study was received from the Ethics Committee for Clinical Trials of Northwestern Tuscany (Study n° 213/201, final approval date: 19/11/2015, Name of the trial: SEMEOTI-CONS, ClinicalTrials.gov Identifier: NCT02818504). Written informed consent was signed by all participants prior to study enrolment in front of a Medical Doctor. All study procedures were designed and conducted in accordance with the tenets of the Declaration of Helsinki.

At baseline, after enrollment all subjects underwent a complete medical history. Table A1 reports the clinical and socio-demographic characteristics of the study population (age, gender, clinical history, lifestyle).

Table A1: Clinical and socio-demographic characteristics of the study population, made of 30 male and 38 female subjects.

		Male	Female	Total	p-value
Age (Mean $\pm$ sd)		46.3 $\pm$ 9.9	45.0 $\pm$ 10.6	45.6 $\pm$ 10.3	0.614
Smoker	No	96.7%	81.6%	88.2%	0.055
	Yes	3.3%	18.4%	11.8%	
Diabetes	No	93.3%	97.4%	95.6%	0.421
	Yes	6.7%	2.6%	4.4%	
Cholesterol	No	66.7%	78.9%	73.5%	0.254
	Yes	33.3%	21.1%	26.5%	
Hypertension	No	90.0%	92.1%	91.2%	0.761
	Yes	10.0%	7.9%	8.8%	

The physical examination consisted of: anthropometric parameters (height, weight, waist and hip circumference); body composition analysis (lean mass and fat mass) by an air displacement pletismograph BodPod (Cosmed, USA); peripheral venous blood samplings (total, HDL and LDL cholesterol, triglycerides, glucose, insulin, HbA1c, haemoglobin, creatinine); AGEs

(Advanced glycation end products) assessed by forearm skin autofluorescence (AGE reader DiagnOptics Technologies, The Netherlands); endothelium-dependent vasodilatation via peripheral arterial tonometry (EndoPAT2000, Itamar Medical Ltd., Caesarea, Israel); heart rate recorded by a standard 12-lead ECG; blood pressure measured non-invasively by a manual sphygmomanometer and averaged over three consecutive measures.

## Appendix B. SOM network

SOMs are unsupervised neural networks having the capability to build accurate, but low-dimensional, topology preserving-maps of the input data space [36]. This means that similar inputs data tend to excite neighboring units in the map. The map space is defined beforehand, usually as a finite two-dimensional region where a set of nodes  $m_i$ ,  $i = 1, \dots, N$  is arranged in a regular grid. Each node is fed by input data  $\mathbf{x}_k$  via a weight vector  $\mathbf{w}_i$ . For a given input  $\mathbf{x}_k$ , the output of the network is defined by the best matching (or winning) unit  $m_c$  obtained by:

$$c = \operatorname{argmin}_k (\|\mathbf{w}_k - \mathbf{x}\|)$$

The weight  $\mathbf{w}_c$  represents the network response and is a point in data space.

During training, nodes in the map space stay fixed, while their weight vectors are moved toward the input data without spoiling the topology induced from the map space. During a training epoch all input patterns are presented to the network. For each pattern, the weight of  $m_c$  unit and neighboring units are adapted according to a predefined neighborhood function  $h_{ck}$  (Gaussian is a common choice for  $h$ ). In this work we adopted the batch version of the SOM adaptation algorithm [36] leading to the adaptation rule:

$$\mathbf{w}_i = \frac{\sum_k h_{ci} \mathbf{x}_k}{\sum_k h_{ci}}$$

This equation ensures a faster convergence and provides more stable results with respect to stochastic adaptation. After training, SOM can build accurate topographic representation of the input space catching significant details including possible data clustering. In particular, each weight vector can be viewed as a prototype in data space as it tends to respond to a set of "near" input points.



625 **References**

- 626 [1] G. A. Roth, et al., Global burden of cardiovascular diseases and  
627 risk factors, 1990–2019: Update from the gbd 2019 study, *Journal of the American College of Cardiology* 76 (25) (2020) 2982–  
628 3021. doi:<https://doi.org/10.1016/j.jacc.2020.11.010>.
- 629 [2] F. Baygi, K. Herttua, O. Jensen, S. Djalalinia, A. M. Ghorabi,  
630 H. Asayesh, M. Qorbani, Global prevalence of cardiometabolic  
631 risk factors in the military population: a systematic re-  
632 view and meta-analysis, *BMC Endocr Disord.* 20 (1) (2020).  
633 doi:[10.1186/s12902-020-0489-6](https://doi.org/10.1186/s12902-020-0489-6).
- 634 [3] R. McQueen, V. Ghushchyan, T. Olufade, J. Sheehan, K. Nair,  
635 J. Saseen, Incremental increases in economic burden parallels  
636 cardiometabolic risk factors in the us, *Diabetes Metab Syndr*  
637 *Obes.* 9 (2016) 233–41. doi:[10.2147/DMSO.S106809](https://doi.org/10.2147/DMSO.S106809).
- 638 [4] K. C. Ferdinand, Global perspectives on cardiometabolic  
639 risk and cardiovascular disease: from basic science to  
640 bedside, *Annals of Translational Medicine* 6 (15) (2018).  
641 doi:[10.21037/atm.2018.07.28](https://doi.org/10.21037/atm.2018.07.28).
- 642 [5] L. Neubeck, G. Coorey, D. Peiris, J. Mulley, E. Hee-  
643 ley, F. Hersch, J. Redfern, Development of an integrated  
644 e-health tool for people with, or at high risk of, cardio-  
645 vascular disease: The consumer navigation of elec-  
646 tronic cardiovascular tools (connect) web application, *Inter-  
647 national Journal of Medical Informatics* 96 (2016) 24–37.  
648 doi:<https://doi.org/10.1016/j.ijmedinf.2016.01.009>.
- 649 [6] L. D. Breeman, M. Keesman, D. E. Atsma, N. H. Cha-  
650 vannes, V. Janssen, L. van Gemert-Pijnen, H. Kemps,  
651 W. Kraaij, F. Rauwers, T. Reijnders, W. Scholte op Reimer,  
652 J. Wentzel, R. A. Kraaijenhagen, A. W. Evers, A multi-  
653 stakeholder approach to ehealth development: Promoting sus-  
654 tained healthy living among cardiovascular patients, *Internat-  
655 ional Journal of Medical Informatics* 147 (2021) 104364.  
656 doi:<https://doi.org/10.1016/j.ijmedinf.2020.104364>.
- 657 [7] R. H. Hoyle, *Handbook of structural equation modeling*, Guil-  
658 ford Press, 2012.
- 659 [8] L. Bastiani, L. Fortunato, S. Pieroni, F. Bianchi, F. Adorni,  
660 F. Prinelli, A. Giacomelli, G. Pagani, S. Maggi, C. Trevisan,  
661 M. Noale, N. Jesuthasan, A. Sojic, C. Pettenati, M. Andreoni,  
662 R. Antonelli Incalzi, M. Galli, S. Molinaro, Rapid covid-19  
663 screening based on self-reported symptoms: Psychometric as-  
664 sessment and validation of the epicovid19 short diagnostic scale,  
665 *J Med Internet Res* 23 (1) (2021) e23897. doi:[10.2196/23](https://doi.org/10.2196/23).
- 666 [9] G. Coppini, R. Favilla, A. Gastaldelli, S. Colantonio,  
667 P. Marraccini, Moving medical semeiotics to the digi-  
668 tal realm - semeiotics approach to face signs of cardio-  
669 metabolic risk, in: *Proceedings of the International Confer-  
670 ence on Health Informatics - Volume 1: SUPERHEAL*,  
671 (BIOSTEC 2014), INSTICC, SciTePress, 2014, pp. 606–613.  
672 doi:[10.5220/0004939106060613](https://doi.org/10.5220/0004939106060613).
- 673 [10] F. Chiarugi, S. Colantonio, D. Emmanouilidou, M. Mar-  
674 tinelli, D. Moroni, O. Salvetti, Decision support in heart  
675 failure through processing of electro- and echocardiograms,  
676 *Artificial Intelligence in Medicine* 50 (2) (2010) 95–104.  
677 doi:[10.1016/j.artmed.2010.05.001](https://doi.org/10.1016/j.artmed.2010.05.001).
- 678 [11] R. Maddison, J. Rawstorn, S. Islam, K. Ball, S. Tighe, N. Gant,  
679 R. M. Whittaker, C. K. Chow, mhealth interventions for exer-  
680 cise and risk factor modification in cardiovascular disease,  
681 *Exercise and sport sciences reviews* 47 (2) (2019) 86–90.  
682 doi:[10.1249/JES.000000000000185](https://doi.org/10.1249/JES.000000000000185).
- 683 [12] A. Dimopoulos, M. Nikolaidou, F. Caballero, W. Engchuan,  
684 A. Sanchez-Niubo, H. Arndt, J. Ayuso-Mateos, J. Haro, S. Chat-  
685 terji, E. Georgousopoulou, C. Pitsavos, D. Panagiotakos, Ma-  
686 chine learning methodologies versus cardiovascular risk scores,  
687 in predicting disease risk, *BMC Med Res Methodol* 18 (1)  
688 (2018) 179. doi:[10.1186/s12874-018-0644-1](https://doi.org/10.1186/s12874-018-0644-1).
- 689 [13] S. Weng, J. Reys, J. Kai, J. Garibaldi, N. Qureshi, Can  
690 machine-learning improve cardiovascular risk prediction us-  
691 ing routine clinical data?, *PLoS ONE* 12 (e0174944) (2017).  
692 doi:[10.1371/journal.pone.0174944](https://doi.org/10.1371/journal.pone.0174944).
- 693 [14] R. M. Conroy, K. Pyörälä, A. Fitzgerald, S. Sans, A. Menotti,  
694 G. D. Bacquer, P. Ducimetière, P. Jousilahti, U. Keil, I. Njølstad,  
695 R. Oganov, T. Thomsen, H. Tunstall-Pedoe, A. Tverdal,  
696 H. Wedel, P. Whincup, L. Wilhelmsen, I. Graham, Score project  
697 group. estimation of ten-year risk of fatal cardiovascular disease  
698 in europe: the score project, *Eur Heart J.* 24 (2003) 987–1003.  
699 doi:[10.1016/s0195-668x\(03\)00114-3](https://doi.org/10.1016/s0195-668x(03)00114-3).
- 700 [15] R. B. D’Agostino, R. Vasan, P. M.J., W. P.A.,  
701 M. Cobain, J. Massaro, W. Kannel, General cardiovas-  
702 cular risk profile for use in primary care: the fram-  
703 ingham heart study, *Circulation* 117 (2008) 743–53.  
704 doi:[10.1161/CIRCULATIONAHA.107.699579](https://doi.org/10.1161/CIRCULATIONAHA.107.699579).
- 705 [16] D. Goff, D. Lloyd-Jones, G. Bennett, S. Coady, R. D’Agostino,  
706 R. Gibbons, P. Greenland, D. Lackland, D. Levy, C. O’Donnell,  
707 J. Robinson, J. Schwartz, S. Shero, S. J. Smith, P. Sor-  
708 lie, N. Stone, P. Wilson, H. Jordan, L. Nevo, J. Wnek,  
709 J. Anderson, J. Halperin, N. Albert, B. Bozkurt, R. Brindis,  
710 L. Curtis, D. DeMets, J. Hocman, J. Hocman, R. Kovacs,  
711 E. Ohman, S. Pressler, F. Sellke, W. Shen, G. Tomaselli,  
712 2013 acc/aha guideline on the assessment of cardiovascu-  
713 lar risk: a report of the american college of cardiol-  
714 ogy/american heart association task force on practice guide-  
715 lines, *Circulation* 129 (2014) S49–73, Erratum in S74–5.  
716 doi:[10.1161/01.cir.0000437741.48606.98](https://doi.org/10.1161/01.cir.0000437741.48606.98).
- 717 [17] K. Hajifathalian, P. Ueda, Y. Lu, M. Woodward, A. Ahmad-  
718 vand, C. Aguilar-Salinas, F. Azizi, R. Cifkova, M. Di Cesare,  
719 L. Eriksen, F. Farzadfar, N. Ikeda, D. Khalili, Y. Khang, V. Lan-  
720 ska, L. León-Muñoz, D. Magliano, K. Msyamboza, K. Oh,  
721 F. Rodríguez-Artalejo, R. Rojas-Martinez, J. Shaw, G. Stevens,  
722 J. Tolstrup, B. Zhou, J. Salomon, M. Ezzati, G. Danaei, A  
723 novel risk score to predict cardiovascular disease risk in na-  
724 tional populations (globoRisk): A pooled analysis of prospective  
725 cohorts and health examination surveys, *The Lancet Diabetes  
726 and Endocrinology* 3 (5) (2015) 339–355. doi:[10.1016/S2213-8587\(15\)00081-9](https://doi.org/10.1016/S2213-8587(15)00081-9).
- 727 [18] G. Bedogni, S. Bellentani, L. Miglioli, M. Passalacqua, A. Cas-  
728 tiglione, C. Tiribelli, The fatty liver index: a simple and accurate  
729 predictor of hepatic steatosis in the general population, *BMC  
730 Gastroenterology* 6 (2006). doi:[10.1186/1471-230X-6-33](https://doi.org/10.1186/1471-230X-6-33).
- 731 [19] J. Lindström, J. Tuomilehto, The diabetes risk score: A practi-  
732 cal tool to predict type 2 diabetes risk, *Diabetes Care* 26 (2003)  
733 725–731. doi:[10.2337/diacare.26.3.725](https://doi.org/10.2337/diacare.26.3.725).
- 734 [20] D. Matthews, J. Hosker, A. S. Rudenski, B. Naylor, D. Treacher,  
735 R. Turner, Homeostasis model assessment: insulin resistance  
736 and beta-cell function from fasting plasma glucose and insulin  
737 concentrations in man, *Diabetologia* 28 (7) (1985) 412–419.  
738 doi:[10.1007/bf00280883](https://doi.org/10.1007/bf00280883).
- 739 [21] K. Karmali, S. D. Persell, P. Perel, D. Lloyd-Jones, M. Berend-  
740 sen, M. Huffman, Risk scoring for the primary prevention of car-  
741 diovascular disease, *Cochrane Database Syst Rev* 3 (3) (2017).  
742 doi:[10.1007/bf00280883](https://doi.org/10.1007/bf00280883).
- 743 [22] R. W. Treskes, E. T. van der Velde, R. Barendse, N. Bru-  
744 ning, Mobile health in cardiology: a review of currently  
745 available medical apps and equipment for remote monitoring,  
746 *Expert Review of Medical Devices* 13 (9) (2016) 823–830.  
747 doi:[10.1080/17434440.2016.1218277](https://doi.org/10.1080/17434440.2016.1218277).
- 748 [23] N. J. Conn, K. Q. Schwarz, D. A. Borkholder, In-home  
749 cardiovascular monitoring system for heart failure: Compar-  
750 ative study, *JMIR Mhealth Uhealth* 7 (1) (2019) e12419.



- 753 doi:10.2196/12419.
- 754 [24] A. D., B. R., A. M., et al., 2019 acc/aha guideline on the primary 819  
755 prevention of cardiovascular disease, *J Am Coll Cardiol* 74 (10) 820  
756 (2019) e177–e232. 821
- 757 [25] M. Khodarahmi, M. Asghari-Jafarabadi, M. Abbasal- 822  
758 izad Farhangi, A structural equation modeling approach 823  
759 for the association of a healthy eating index with 824  
760 metabolic syndrome and cardio-metabolic risk factors 825  
761 among obese individuals, *PLOS ONE* 14 (7) (2019) 1–20. 826  
762 doi:10.1371/journal.pone.0219193. 827  
763 URL <https://doi.org/10.1371/journal.pone.0219193> 828
- 764 [26] L. L. R. Rodrigues, D. K. Shetty, N. Naik, C. B. Maddodi, 829  
765 A. Rao, A. K. Shetty, R. Bhat, B. M. Z. Hameed, Machine learn- 830  
766 ing in coronary heart disease prediction: Structural equation 831  
767 modelling approach, *Cogent Engineering* 7 (1) (2020) 1723198. 832  
768 doi:10.1080/23311916.2020.1723198. 833
- 769 [27] N. Shakibaei, R. Hassannejad, N. Mohammadifard, H. R. 834  
770 Marateb, M. Mansourian, M. A. Mañanas, N. Sarrafzadegan, 835  
771 Pathways leading to prevention of fatal and non-fatal cardio- 836  
772 vascular disease: An interaction model on 15 years population- 837  
773 based cohort study, *Lipids in Health and Disease* 19 (1) (2020) 838  
774 203. doi:10.1186/s12944-020-01375-8.
- 775 [28] P. O. Bonetti, G. M. Pumper, S. T. Higano, D. R. Holmes, J. T. 839  
776 Kuvin, A. Lerman, Noninvasive identification of patients with 840  
777 early coronary atherosclerosis by assessment of digital reac- 841  
778 tive hyperemia, *Journal of the American College of Cardiology* 842  
779 44 (11) (2004) 2137–2141. doi:10.1016/j.jacc.2004.08.062.
- 780 [29] P. Henriquez, B. J. Matuszewski, Y. Andreu-Cabedo, L. Bas- 843  
781 tiani, S. Colantonio, G. Coppini, M. D’Acunto, R. Favilla, 844  
782 D. Germanese, D. Giorgi, P. Marraccini, M. Martinelli, M.- 845  
783 A. Morales, M. A. Pascali, M. Righi, O. Salvetti, M. Lars- 846  
784 son, T. Strömberg, L. Randeberg, A. Bjorgan, G. Giannakakis, 847  
785 M. Pediaditis, F. Chiarugi, E. Christinaki, K. Marias, M. Tsik- 848  
786 nakis, Mirror mirror on the wall... an unobtrusive intelligent 849  
787 multisensory mirror for well-being status self-assessment and 850  
788 visualization, *IEEE Transactions on Multimedia* 19 (7) (2017) 851  
789 1467–1481. doi:10.1109/TMM.2017.2666545.
- 790 [30] M. Pascali, D. Giorgi, L. Bastiani, E. Buzzigoli, P. Henriquez, 852  
791 B. Matuszewski, M.-A. Morales, S. Colantonio, Face mor- 853  
792 phology: Can it tell us something about body weight and 854  
793 fat?. *Computers in Biology and Medicine* 76 (2016) 238–249. 855  
794 doi:10.1016/j.compbiomed.2016.06.006.
- 795 [31] R. Singh, A. Barden, T. Mori, L. Beilin, Advanced glycation 856  
796 end-products: a review, *Diabetologia* 44 (2) (2001) 129–146.
- 797 [32] R. Meerwaldt, R. Graaff, P. H. N. Oomen, T. P. Links, J. J. 857  
798 Jager, N. L. Alderson, S. R. Thorpe, J. W. Baynes, R. O. B. 858  
799 Gans, A. J. Smit, Simple non-invasive assessment of advanced 859  
800 glycation endproduct accumulation, *Diabetologia* 47 (7) (2004) 860  
801 1324–1330.
- 802 [33] M. Larsson, R. Favilla, T. Strömberg, Assessment of advanced 861  
803 glycated end product accumulation in skin using auto fluores- 862  
804 cence multispectral imaging, *Comput Biol Med* 85 (2017) 106– 863  
805 111. doi:10.1016/j.compbiomed.2016.04.005.
- 806 [34] M. J. Joyner, N. M. Dietz, J. T. Shepherd, From belfast to 864  
807 mayo and beyond: the use and future of plethysmography to 865  
808 study blood flow in human limbs, *Journal of Applied Physiology* 866  
809 91 (6) (2001) 2431–2441. doi:10.1152/jappl.2001.91.6.2431.
- 810 [35] S. Bergstrand, M.-A. Morales, G. Coppini, M. Larsson, 867  
811 T. Strömberg, The relationship between forearm skin speed- 868  
812 resolved perfusion and oxygen saturation, and finger arter- 869  
813 ial pulsation amplitudes, as indirect measures of endo- 870  
814 thelial function, *Microcirculation* 25 (2) (2018) e12422. 871  
815 doi:<https://doi.org/10.1111/micc.12422>.
- 816 [36] T. Kohonen, *Self-Organizing Maps*, Springer, Berlin, 2001. 872  
817 doi:10.1007/978-3-642-56927-2.
- 818 [37] I. D., Structural equations modeling: Fit indices, sample size, 873  
819 and advanced topics, *Journal of Consumer Psychology* 20 874  
820 (2010) 90–98.
- 821 [38] J. Vesanto, E. Alhoniemi, Clustering of the self-organizing map, 875  
822 *IEEE Transactions on Neural Networks* 11 (3) (2000) 586–600. 876  
823 doi:10.1109/72.846731.
- 824 [39] K. H. C. Li, F. A. White, T. Tipoe, T. Liu, M. C. Wong, A. Je- 877  
825 suthasan, A. Baranchuk, G. Tse, B. P. Yan, The current state 878  
826 of mobile phone apps for monitoring heart rate, heart rate vari- 879  
827 ability, and atrial fibrillation: Narrative review, *JMIR Mhealth 880*  
828 *Uhealth* 7 (2) (2019). doi:10.2196/11606. 881  
829 URL <http://mhealth.jmir.org/2019/2/e11606/>
- 830 [40] R. Li, N. Liang, F. Bu, T. Hesketh, The effectiveness of self- 882  
831 management of hypertension in adults using mobile health: Sys- 883  
832 tematic review and meta-analysis, *JMIR Mhealth Uhealth* 8 884  
833 (2020) e17776. doi:10.2196/17776.
- 834 [41] H. Darvishi, D. Ciunzo, E. R. Eide, P. S. Rossi, Sensor-fault 885  
835 detection, isolation and accommodation for digital twins via 886  
836 modular data-driven architecture, *IEEE Sensors Journal* 21 (4) 887  
837 (2021) 4827–4838. doi:10.1109/JSEN.2020.3029459.

Manuscript Title:

**Cardio-metabolic risk modeling and assessment through sensor-based measurements**

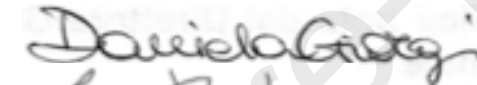
Authors:

*Daniela Giorgi, Luca Bastiani, Maria Aurora Morales, Maria Antonietta Pascali, Sara Colantonio, Giuseppe Coppini*

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Signed by all authors as follows:

Daniela Giorgi



Luca Bastiani



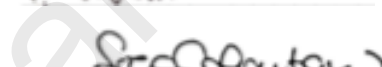
Maria Aurora Morales



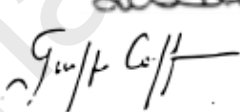
Maria Antonietta Pascali



Sara Colantonio



Giuseppe Coppini



Date 07/02/2022

Manuscript Title:

**Cardio-metabolic risk modeling and assessment through sensor-based measurements**

Authors:

*Daniela Giorgi, Luca Bastiani, Maria Aurora Morales, Maria Antonietta Pascali, Sara Colantonio, Giuseppe Coppini*

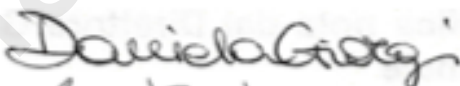
We the undersigned declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We understand that the Corresponding Author is the sole contact for the Editorial process. He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs

Signed by all authors as follows:

Daniela Giorgi



Luca Bastiani



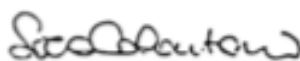
Maria Aurora Morales



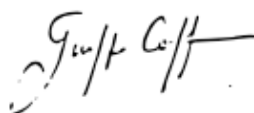
Maria Antonietta Pascali



Sara Colantonio



Giuseppe Coppini



Date 07/02/2022