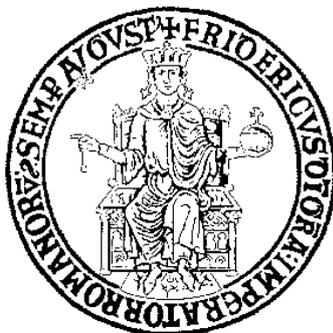


UNIVERSITÀ DEGLI STUDI DI NAPOLI "FEDERICO II"  
DIPARTIMENTO DI MEDICINA CLINICA E CHIRURGICA



DOTTORATO DI RICERCA IN  
TERAPIE AVANZATE BIOMEDICHE E CHIRURGICHE  
*XXXI° CICLO*

**TESI DI DOTTORATO:**

**Ruolo di uno score di stato d'infiammazione  
nell'aterosclerosi subclinica analizzati mediante  
metodologie statistiche innovative su dati di uno  
studio longitudinale europeo.**

**Coordinatore del dottorato e Tutor**

*Ch.mo Prof. Di Minno Giovanni*

**Dottorando**

*Alice Bonomi*

**ANNO ACCADEMICO 2017-2018**

# Sommario

<b>RIASSUNTO</b> .....	<b>3</b>
<b>ABSTRACT</b> .....	<b>5</b>
<b>INTRODUZIONE</b> .....	<b>7</b>
<b>ANALISI STATISTICA IN ARTICOLI BIOMEDICI</b> .....	<b>7</b>
<b>MODELLI A REGRESSIONE MULTIPLA</b> .....	<b>11</b>
<b>MODELLI AD EQUAZIONI STRUTTURALI E PATH ANALYSIS</b> .....	<b>13</b>
<i>Introduzione ai SEM e Causalità</i> .....	<b>13</b>
<i>Storia dei SEM e della Path Analysis</i> .....	<b>18</b>
<i>Path Analysis</i> .....	<b>21</b>
<i>SEM</i> .....	<b>23</b>
<b>RISCHIO CARDIOVASCOLARE, ATEROSCLEROSI ED INFIAMMAZIONE</b> .....	<b>27</b>
<b>MARKERS DELL'ATEROSCLEROSI: SPESSORE DEL COMPLESSO MEDIO-INTIMALE (IMT)</b> .....	<b>32</b>
<b>EVENTI VASCOLARI MAGGIORI</b> .....	<b>37</b>
<i>Malattia cardiovascolare (CVD)</i> .....	<b>37</b>
<i>Malattia cerebrovascolare</i> .....	<b>37</b>
<i>Malattie vascolari periferiche</i> .....	<b>38</b>
<b>RAZIONALE</b> .....	<b>41</b>
<b>OBIETTIVO DELLO STUDIO</b> .....	<b>43</b>
<b>METODI</b> .....	<b>44</b>
<b>POPOLAZIONE</b> .....	<b>44</b>
<i>Lo studio IMPROVE</i> .....	<b>44</b>
<b>MISURA DELLO SPESSORE DEL COMPLESSO MEDIO-INTIMALE (IMT) E SUA VARIAZIONE NEL TEMPO</b> .....	<b>46</b>
<b>ANALISI DI LABORATORIO</b> .....	<b>49</b>
<b>ANALISI STATISTICA</b> .....	<b>51</b>
<i>Score di stato infiammatorio</i> .....	<b>52</b>
<i>Path Analysis e SEM</i> .....	<b>52</b>
<i>Indici di adattamento</i> .....	<b>53</b>
<b>RISULTATI</b> .....	<b>55</b>
<b>CARATTERISTICHE DELLA POPOLAZIONE</b> .....	<b>55</b>
<b>RELAZIONE C-IMT E BIOMARCATORI DELLO STATO INFIAMMATORIO</b> .....	<b>58</b>
<i>Correlazione reciproca dei Biomarcatori dello stato infiammatorio</i> .....	<b>58</b>
<i>Analisi con modelli di regressione semplice e multipla della relazione tra i Biomarcatori dello stato infiammatorio e c-IMT</i> .....	<b>59</b>
<b>ASSOCIAZIONE EVENTI E BIOMARCATORI DELLO STATO INFIAMMATORIO: MODELLI DI REGRESSIONE</b> .....	<b>60</b>
<b>FATTORI DI RISCHIO, C-IMT, EVENTI E SCORE DELLO STATO INFIAMMATORIO</b> .....	<b>61</b>
<i>Relazione tra score di stato di infiammazione e fattori di rischio, c-IMT</i> .....	<b>61</b>
<i>Capacità predittiva degli eventi da parte dello score di stato di infiammazione: confronto con i singoli biomarcatori</i> .....	<b>62</b>
<i>Relazione tra score di stato di infiammazione ed eventi vascolari maggiori</i> .....	<b>64</b>
<b>ASSOCIAZIONE TRA EVENTI, C-IMT E FATTORI DI RISCHIO: PATH ANALYSIS E SEM</b> .....	<b>66</b>
<i>Associazione eventi, c-IMT e fattori di rischio</i> .....	<b>66</b>
<i>Associazione tra eventi, progressione del c-IMT e fattori di rischio</i> .....	<b>69</b>
<i>Associazione tra eventi, c-IMT, fattori di rischio e score dello stato infiammatorio</i> .....	<b>72</b>
<i>Associazione tra eventi, progressione c-IMT, fattori di rischio e score di stato infiammatorio</i> .....	<b>76</b>
<i>Associazione tra eventi, c-IMT, progressione c-IMT, fattori di rischio e score di stato infiammatorio</i> .....	<b>80</b>
<i>Associazione tra eventi, c-IMT, progressione c-IMT, fattori di rischio e variabile latente 'Infiammazione'</i> .....	<b>83</b>

<b>DISCUSSIONE .....</b>	<b>86</b>
<b>RELAZIONI REGRESSIONI VS RELAZIONI STIMATE CON PATH ANALYSIS E SEM. ....</b>	<b>87</b>
<b>SCORE DI STATO DI INFIAMMAZIONE, C-IMT, PROGRESSIONE C-IMT ED EVENTI VASCOLARI MAGGIORI. ....</b>	<b>89</b>
<b>CONFRONTO TRA LO SCORE DI STATO DI INFIAMMAZIONE E LA VARIABILE LATENTE 'INFIAMMAZIONE' .....</b>	<b>90</b>
<b>LIMITAZIONI .....</b>	<b>92</b>
<b>CONCLUSIONI E RILEVANZA .....</b>	<b>93</b>
<b>REFERENZE.....</b>	<b>95</b>

## Riassunto

**STATO DELL'ARTE.** La statistica medica, come altre discipline della scienza, è un settore in continua evoluzione, dove metodi nuovi e complessi vengono continuamente sviluppati e incorporati in diverse discipline. Ad oggi, digitando 'regression' su PubMed troviamo circa 700.000 risultati; dettagliando la ricerca utilizzando 'multiple regression' individuiamo circa 130.000 articoli. Invece, digitando metodologie innovative come 'Path analysis' o 'structural equation modeling', gli articoli identificati sono circa 20.000 e 9.000, rispettivamente. Questi dati mostrano come nella ricerca medica sia molto più frequente usare metodologie statistiche molto consolidate, come regressioni semplici o multivariate, piuttosto che metodologie innovative più sofisticate, anche se negli ultimi anni il loro utilizzo è aumentato considerevolmente.

**OBIETTIVI E DATI ANALIZZATI.** Abbiamo utilizzato i dati dello studio IMPROVE - uno studio multicentrico, longitudinale, osservazionale, svolto in cinque paesi europei, con lo scopo di valutare la relazione tra uno score che rifletta lo 'stato infiammatorio' dei soggetti e lo spessore medio-intimale della carotide (c-IMT) al basale, la sua progressione nel tempo e gli eventi vascolari. Abbiamo confrontato metodologie statistiche innovative, quali la Path-Analysis ed i modelli ad equazioni strutturali (SEM), con metodologie tradizionali, quali regressione multipla, regressione logistica e modello di Cox.

**METODI.** Lo score di 'stato infiammatorio' è stato creato utilizzando un pannello di biomarcatori di infiammazione, combinati con pesi proporzionali alla loro associazione multivariata con l'aterosclerosi subclinica. Per studiare le possibili associazioni tra i fattori di rischio, le caratteristiche cliniche dei soggetti, c-IMT, la progressione e gli eventi vascolari, e lo score è stata implementata una Path Analysis e i SEM, e i risultati sono stati rappresentati mediante grafici che evidenziano le relazioni dirette tra le variabili.

**RISULTATI.** Il database utilizzato comprendeva 3703 soggetti, arruolati in 7 differenti centri: Kuopio (2 centri), Stoccolma, Groningen, Parigi, Milano, Perugia. Al momento dell'arruolamento i pazienti erano esenti da eventi cardio/cerebro cardiovascolari, ma

presentavano almeno tre fattori di rischio vascolare. Il 48% dei soggetti erano maschi e avevano un'età media di  $64,2 \pm 5.4$  anni.

Dopo un tempo mediano di follow-up di 36.2 mesi [IQ range:35.8-37.4], si sono verificati 215 eventi vascolari (125 Cardiovascolari, 73 Cerebro-cardiovascolari, 17 periferici). Implementando la Path Analysis, abbiamo dimostrato una relazione positiva e molto significativa tra score di 'stato infiammatorio' e c-IMT basale; inoltre, lo score era correlato positivamente agli eventi vascolari, ma non alla progressione del c-IMT.

**CONCLUSIONI E RILEVANZA.** I nostri risultati indicano che i metodi statistici innovativi, quali la Path Analysis e SEM, possono essere molto utili nelle ricerche biomediche, infatti, hanno mostrato tutte le relazioni che eravamo interessati a studiare in un unico grafico e hanno identificato gli effetti indiretti, che non siamo in grado di valutare con i modelli di regressione multipla 'classici'. I nostri risultati, inoltre, indicano che l'infiammazione, misurata con uno score derivato da un pannello di variabili misurate nel sangue, è fortemente associata al livello aterosclerosi subclinica presente nell'organismo (con una relazione che potrebbe essere di causa, ma anche di effetto); inoltre, lo score di stato infiammatorio identificato predice gli eventi vascolari indipendentemente dai classici fattori di rischio. Tale risultato indica l'opportunità di misurare diversi biomarcatori di infiammazione al fine di perfezionare la stima del rischio individuale.

# **Abstract**

**BACKGROUND** Medical statistics, like other disciplines of science, is a constantly evolving field, where new and complex methods are continuously developed and incorporated into different disciplines. To date, typing 'regression' on PubMed we find about 700,000 results; by detailing the search using 'multiple regression' we obtain about 130,000 articles. Instead, by searching innovative methodologies such as 'Path analysis' or 'structural equation modeling' we will identify about 20,000 and 9,000 articles, respectively. These data show that, in medical research, well established statistical methodologies, such as simple or multivariate regressions, are used much more commonly than more sophisticated, innovative methodologies, although the use of the latter has increased considerably in recent years.

**OBJECTIVES AND EVIDENCE ANALYZED.** We used data from the IMPROVE study - a multicenter, longitudinal, observational study, carried out in five European countries - with the aim of assessing the relationship between a score that reflects the 'inflammatory state' of the subjects and the intima-media thickness of the carotid artery (c-IMT) at baseline, its change over time and the incidence of vascular events. Innovative statistical methodologies, such as Path-Analysis and structural equations modeling (SEM), were compared with traditional ones, like multiple linear regression, logistic regression and Cox model.

**METHODS.** The 'inflammatory state' score was created by using a panel of biomarkers of inflammation, combined with weights proportional to their multivariate association with subclinical atherosclerosis. To study the possible associations between risk factors, clinical characteristics of subjects, c-IMT- progression and vascular events, and the score, a Path Analysis and SEMs were implemented. Results were represented by graphs highlighting the direct relationships between variables.

**RESULTS.** The database included 3703 subjects, enrolled in 7 different centers: Kuopio (2 centers), Stockholm, Groningen, Paris, Milan, Perugia. At the time of enrolment, patients had no history of vascular events, but had at least three vascular risk factors. 48% of the subjects were male and the average age was  $64.2 \pm 5.4$  years. After a median follow-up time of 36.2 months [IQ range:35.8-37.4], 215 vascular events were

identified (125 Cardiovascular, 73 Cerebro-cardiovascular, 17 peripheral). By implementing the Path Analysis, a positive and significant relationship between 'inflammatory state' score and basal c-IMT was highlighted; moreover, the score was positively correlated to vascular events, but not to c-IMT progression.

**CONCLUSIONS AND RELEVANCE.** Our results indicate that innovative statistical methods, such as Path Analysis and SEM, can be very useful in biomedical research, in fact, they disclosed all the relevant relationships in a single graph and identified also the indirect effects, which cannot be evaluated with 'classical' regression models. Our results also indicate that inflammation, measured with a score derived from a panel of variables measured in the blood, is strongly associated with the level of subclinical atherosclerosis (with a relationship that could be of cause, but also of effect); in addition, the inflammatory state score predicts vascular events independently from classical risk factors. This result indicates the opportunity to measure different biomarkers of inflammation in order to improve the estimation of individual risk. Needless to say, further studies are needed in order to validate the score and to make it usable in clinical practice.

# Introduzione

## Analisi statistica in articoli biomedici

La statistica medica, come altre discipline della scienza, è un settore in continua evoluzione, dove metodi nuovi e complessi vengono continuamente sviluppati e incorporati in diverse discipline per aiutare a risolvere problemi. Molti metodi sempre più sofisticati vengono implementati, grazie anche alla disponibilità di pacchetti di software statistici multidimensionali, ma i metodi 'semplici', come i test t o chi-quadro, rimangono quelli più utilizzati. (Strasak, Zaman et al. 2007)

L'uso della statistica nelle riviste mediche è oggetto di notevoli discussioni da diversi anni e vi è un ampio consenso sul fatto che spesso vengono preferite metodologie statistiche meno sofisticate, verosimilmente per trovare un punto d'incontro tra medici/biologici e statistici/matematici. (Altman 1981; Altman and Goodman 1994) .

Nel luglio 2003, Taback e Krzyzanowska hanno condotto una revisione dell'uso di metodi statistici pubblicati negli abstract di quattro riviste mediche ad alto impatto (NEJM, BMJ, JAMA, Lancet) ed è emerso che i metodi più comunemente usati erano statistiche riassuntive descrittive insieme alle classiche procedure di regressione. (Taback and Krzyzanowska 2008)

Nel 2013 Nietert et al. hanno riscontrato che di 127.469 pubblicazioni del General Internal Medicine (GIM) solo l'1.7% hanno citato articoli di statistica. Inoltre, hanno analizzato quali erano gli argomenti statistici più citati (Tabella X), concludendo che gli sviluppi nella metodologia statistica non vanno di pari passo con il loro utilizzo nelle ricerche biomediche. (Nietert, Wahlquist et al. 2013)

Nel 2016, Christopher Meaney et al. hanno condotto uno studio di Text mining per descrivere le tendenze nell'uso di metodi statistici ed epidemiologici nella letteratura medica negli ultimi due decenni, per un totale di 113.450 articoli. Mentre le statistiche descrittive numeriche erano utilizzate nel 96,4% degli articoli, solo 23,2% citava l'ANOVA ed il 22.6% la regressione. (Meaney, Moineddin et al. 2016)

Ad oggi, digitando 'regression' su Pubmed troviamo 693.674 risultati, dettagliando la ricerca utilizzando 'multiple regression' otteniamo 128.763 articoli. Invece, digitando metodologie innovative come 'Path analysis' o 'structural equation modeling' vengono identificati 21.247 e 9.318 articoli, rispettivamente (Figura 1). Questi dati mostrano come nella ricerca medica sia molto più frequente usare metodologie statistiche molto consolidate, come regressioni semplici o multivariate, piuttosto che metodologie innovative più sofisticate, anche se negli ultimi anni il loro utilizzo è aumentato considerevolmente.

Table V. Keyword mapping categories and citation frequencies.		
Keyword mapping category	Common individual keywords within mapping category	Number of times keyword mapping category was cited
General/other statistical	models, statistical; data interpretation, statistical; biometry	4806
Randomized trials	randomized controlled trials as topic; clinical trials as topic; treatment outcome	2003
Epidemiology	risk factors; risk; epidemiologic methods	1463
Meta-analysis	meta-analysis; metaregression; inverse probability weighted estimation	1223
Generalized linear models	analysis of variance; regression analysis; logistic models	1043
Computer simulation	computer simulation; Monte Carlo method; Markov chains; bootstrap	619
Analysis of correlated data	cluster analysis; longitudinal studies; mixed models	610
Missing data	missing data; multiple imputation; missing at random	529
Bias	bias (epidemiology); publication bias; file drawer problem	451
Survival analysis	survival analysis; proportional hazards models; survival-time	435
Sample size/power estimation	sample size; power; statistical power	426
Diagnostic testing	sensitivity and specificity; ROC curve; predictive value of tests	355
Bayesian methods	Bayes' theorem; Winbugs; empirical Bayes	349
Study design	research design; design; 2-period cross-over design	333
Bioinformatics methods	models, biological; oligonucleotide array sequence analysis; gene expression profiling	298
Statistical software	Software; st0067 (a Stata program for multiple imputation); databases, factual	194
Econometrics	cost-benefit analysis; cost analysis; costs and cost analysis	192
Non/Semi-parametrics	statistics, nonparametric; semiparametrics; nonparametric estimation	156
Reproducibility of results	reproducibility of results	104
Psychometrics	psychometrics; kappa statistic	64
Multiple comparisons	Bonferroni procedure; false discovery rate; multiple comparisons procedures	57
Non-statistical terminology	humans; female; male	16,097

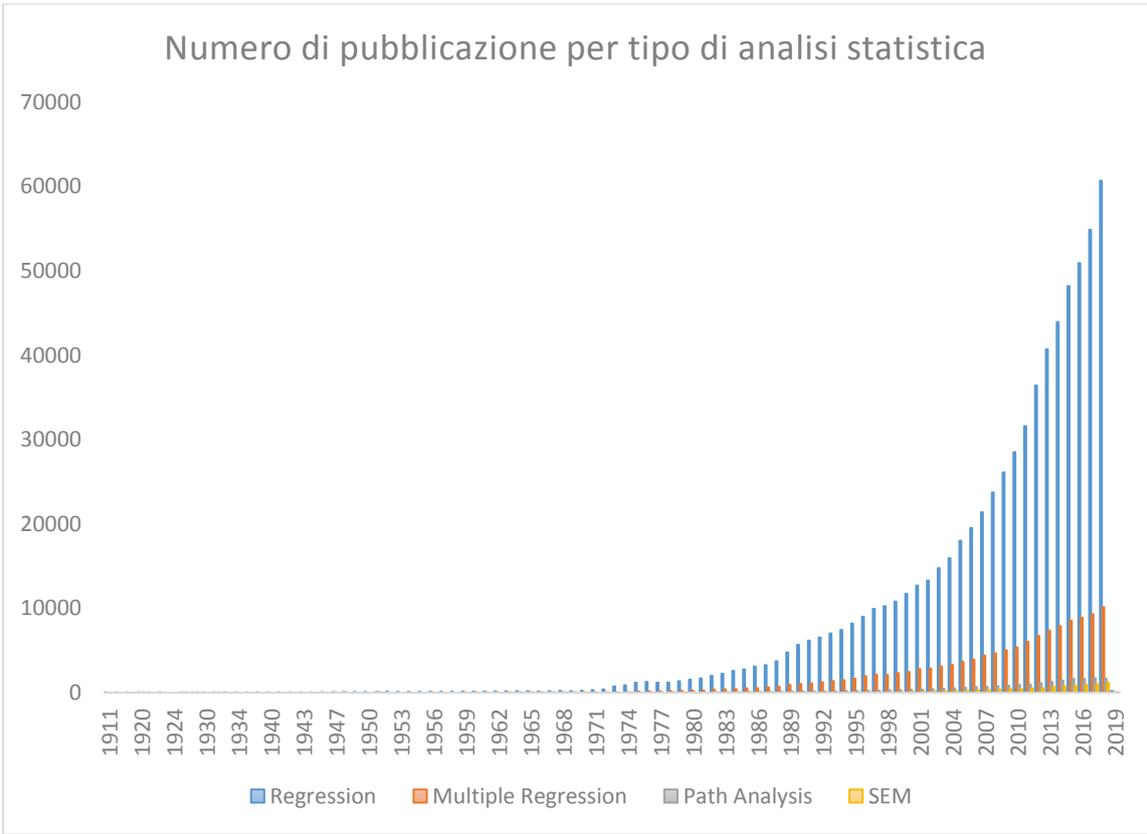


Figura 1: Fonti di dati - PubMed

## Modelli a Regressione Multipla

La regressione multipla è un metodo statistico implementato per analizzare le relazioni tra una variabile risposta Y e diverse variabili indipendenti X. È uno dei metodi statistici più utilizzati nella ricerca biomedica. (Yergens, Dutton et al. 2014)

Considerando il caso più semplice di due variabili indipendenti ed una regressione di tipo lineare, l'equazione è la seguente:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2$$

Il coefficiente di regressione multipla  $\beta_1$  misura la variazione media di Y quando  $X_1$  varia di una sola unità, e  $X_2$  è tenuta costante; lo stesso vale per il coefficiente  $X_2$ . I coefficienti  $\beta_i$ , quindi, sono anche chiamati coefficienti di regressione parziale (o 'netti'). Questa definizione rimarca la differenza nei confronti del coefficiente di regressione semplice lineare, dove  $\beta$  viene indicato come coefficiente di regressione totale.

I modelli di regressione multipla hanno un duplice scopo:

- Esplicativo: studiare e valutare gli effetti delle variabili indipendenti sulla variabile dipendente in funzione di un determinato modello teorico
- Predittivo: individuare una combinazione lineare di variabili indipendenti per predire in modo ottimale il valore assunto dalla variabile dipendente.

Alcuni passi fondamentali nella regressione multipla sono:

- 1) Valutazione dell'adeguatezza delle variabili (livelli di misura, collinearità tra i predittori).
- 2) Scelta della strategia analitica per inserire le variabili indipendenti. Ad esempio, non si devono escludere variabili rilevanti ed introdurre variabili prive di senso.
- 3) Interpretazione della soluzione.

#### 4) Verifica della capacità predittiva delle variabili.

Una regressione multipla può essere implementata in un gran numero di modi diversi.

Le principali strategie sono 3:

1. Standard: tutte le variabili indipendenti vengono inserite insieme nel modello, ed ogni variabile indipendente viene valutata in termini di cosa aggiunge alla predizione della variabile dipendente rispetto alla predizione delle altre variabili indipendenti.
2. Gerarchica: le variabili indipendenti sono inserite nell'equazione in un ordine specifico deciso dal ricercatore, singolarmente o a blocchi. Ogni variabile o blocco di variabile indipendenti viene valutato in termini di cosa aggiunge alla spiegazione della variabile dipendente al momento del suo ingresso.
3. Statistica: è una procedura in cui l'ordine d'entrata delle variabili nel modello è basato unicamente su modelli statistici (ad esempio, selezione *stepwise*). (Hocking 1976)

# Modelli Ad Equazioni Strutturali e Path Analysis

## Introduzione ai SEM e Causalità

I modelli di equazioni strutturali (SEM) sono una tecnica di analisi statistica multivariata che permette di verificare simultaneamente l'influenza di un insieme di variabili su altre. (Stein, Morris et al. 2017)

Possiamo per ciò definire un modello ad equazioni strutturali un modello statistico nel quale ogni equazione rappresenta un legame causale, piuttosto che una mera associazione empirica. (Goldberger 1972)

Si utilizza questo tipo di modello quando si è intenzionati a:

- studiare il fenomeno in esame in termini di causa/effetto;
- implementare analisi che comprendono variabili non misurabili; infatti i SEM sono in grado di studiare le relazioni che intercorrono tra questi tipi di variabili, dette latenti. Ad esempio, nelle scienze sociali, l'intelligenza e la personalità sono costrutti non direttamente misurabili. In campo medico, 'fragilità' o 'stato infiammatorio' potrebbero essere considerate variabili latenti.

Altri tipi di modelli che possiamo considerare parte di questa famiglia sono la *path analysis*, l'analisi fattoriale esplorativa, l'analisi fattoriale confermativa e i modelli di equazioni simultanee. Essi sono in grado di analizzare sia modelli semplici che modelli di difficile o impossibile trattazione con i classici modelli di regressione.

Facendo un confronto tra un modello lineare multivariato e un modello strutturale si può chiarire l'utilità dell'implementazione di quest'ultimo per alcuni aspetti.

Supponendo di volere studiare il percorso scuola-lavoro (Y):

$y_1$ : voto laurea

$y_2$ : reddito primo lavoro

$y_3$ : reddito attuale

avendo a disposizione le seguenti informazioni:

$X_1$ : genere,  
 $X_2$ : reddito genitori,  
 $X_3$ : IQ,  
 $X_4$ : nazionalità.

Utilizzando un modello lineare:

$$\begin{aligned} Y_1 &= c_{11} X_1 + c_{12} X_2 + c_{13} X_3 + c_{14} X_4 + g_1 \\ Y_2 &= c_{21} X_1 + c_{22} X_2 + c_{23} X_3 + c_{24} X_4 + g_2 \\ Y_3 &= c_{31} X_1 + c_{32} X_2 + c_{33} X_3 + c_{34} X_4 + g_3 \end{aligned}$$

Questa equazione ha senso da un punto di vista statistico, ma non permette di considerare eventuali relazioni di causa effetto nel percorso scuola-lavoro (ovvero tra le variabili  $Y_1, Y_2$  e  $Y_3$ ), come invece è possibile fare attraverso un modello ad equazioni strutturali come questo:

$$\begin{aligned} Y_1 &= c_{11} X_1 + c_{12} X_2 + c_{13} X_3 + c_{14} X_4 + g_1 \\ Y_2 &= p_{21} Y_1 + c_{21} X_1 + c_{22} X_2 + c_{23} X_3 + c_{24} X_4 + g_2 \\ Y_3 &= p_{31} Y_1 + p_{32} Y_2 + p_{31} X_1 + c_{32} X_2 + c_{33} X_3 + c_{34} X_4 + g_3. \end{aligned}$$

In questo caso, infatti, è stata introdotta la dinamica dell'interazione tra variabili  $y_j$  (mediante i termini  $p_{21}Y_1, p_{31}Y_1$ , e  $p_{31}Y_2$ ) che permette di tenere conto dei nessi causali tra di esse; ad esempio è possibile valutare la dipendenza tra il reddito al primo lavoro ( $Y_2$ ) e il voto di laurea ( $Y_1$ ).

Detto ciò, è necessario però evidenziare cosa si intende per *causalità*, facendo notare le differenze rispetto al significato di *covariazione*. Si parla di *covariazione* (covarianza, associazione o correlazione) quando si osservano due variabili che presentano

variazioni concomitanti: al variare di una si osserva anche il variare dell'altra.  
(Anderson 1958)

Quando si parla, invece, di *causazione* si intende il concetto di “*produzione*”, ovvero quando la variazione di una variabile *produce* la variazione di un'altra.

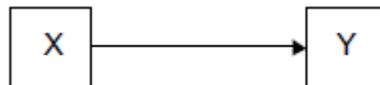
Per meglio distinguere le due terminologie, bisogna tenere in considerazione che la causazione, al contrario della covariazione, presenta:

1. La ***direzionalità*** dell'azione: esiste nella causazione una causa-effetto, ovvero come detto sopra, la variazione di una variabile precede il variare dell'altra, mentre nella covariazione questo non avviene, in quanto le variazioni possono avvenire contemporaneamente.
2. Il ***legame diretto*** tra due variabili: il variare dell'una è *dovuto* (e non solo *associato*) al variare dell'altra.

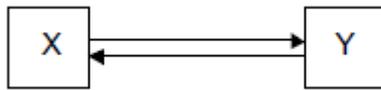
Un'altra importante differenza, da sottolineare è che la covariazione è empiricamente osservabile, mentre la causazione è piuttosto teorica.

Si possono classificare 5 tipi di relazioni casuali:

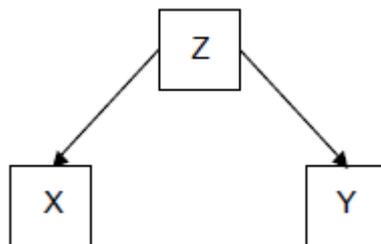
1. Relazione causale *diretta*: vi è questo tipo di relazione quando un mutamento nella variabile causa (X) produce un mutamento nella variabile effetto (Y):



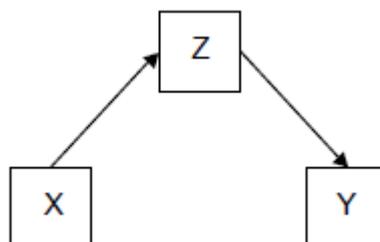
2. Relazione causale *reciproca*: in questa situazione non vi è una situazione causa-effetto lineare, in quanto le variabili si influenzano reciprocamente:



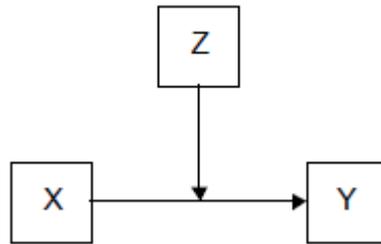
3. Relazione causale *spuria*: è la situazione in cui vi è una situazione di covariazione ma non di causazione. Si nota un covariazione tra due variabili X e Y, provocata però da una terza variabile Z che agisce in maniera causale sulle prime due:



4. Relazione causale *indiretta*: come visto nel punto precedente vi è una terza variabile che media il legame di altre due. In questo caso però la variabile Z funge da ponte tra X e Y che, risultano causalmente collegate attraverso Z:



5. Relazione causale *condizionata*: la relazione tra due variabili (X,Y) cambia a seconda del valore assunto da una terza variabile Z:



## Storia dei SEM e della Path Analysis

Regressione, analisi fattoriale SEM e Path Analysis si sono sviluppati in tempi diversi.

Nel 1896 Karl Pearson creò una formula per la valutazione della correlazione tra due variabili. Ciò gli permise di sviluppare un modello di regressione lineare con l'uso di coefficienti di correlazione, calcolando i pesi di regressione attraverso il metodo dei minimi quadrati.

Alcuni anni dopo, Charles Spearman (1904, 1927) fu il primo, utilizzando questi coefficienti di correlazione, a parlare di analisi fattoriale. La sua idea si basava sulla seguente ipotesi: se gli item di un set sono correlati tra loro, allora le variabili individuali del set di item possono essere combinate per ottenere un punteggio che misura e definisce un *pattern*.

Nel 1955 Howe e, successivamente, nel 1956 Anderson e Rubin e Lawley nel 1958, furono i primi ad utilizzare il termine '*analisi fattoriale confermativa*', che però venne sviluppato solamente qualche anno dopo da Joreskog con il primo articolo nel 1969 e con la creazione di un software ad hoc. (Warr, Lee et al. 1969)

Sewall Wright, genetista, nel 1918 fu invece il primo studioso a parlare di *Path Analysis*. Pubblicò la prima applicazione di essa, modellando la misura ossea dei conigli. (Wright 1918). Egli sviluppò, attraverso l'utilizzo di coefficienti di correlazione e dell'analisi di regressione, un modello più complesso per mettere in relazione tutte le variabili osservate. Inizialmente aveva calcolato tutte le correlazioni parziali tra le misure, senza però rimanere soddisfatto dei risultati ottenuti, in quanto non permettevano di esplicitare eventuali nessi causali. Allora, impose una struttura causale alle correlazioni osservate, per mezzo coefficienti definiti *strutturali*, scomponendo la varianza delle osservazioni dovuta a varie cause ereditarie, e sviluppando così la path analysis.

Successivamente, applicò questo tipo di analisi a sistemi di incroci tra le cavie, ponendo le basi per gran parte della successiva genetica delle popolazioni. Ad esempio, ha modellato la proporzione di bianco nelle cavie maculate, scomponendo la varianza in ereditarietà ( $h$ ), in ambiente comune della lettiera ( $e$ ) ed in altri fattori, come il “*developmental noise*” (ovvero le differenze di fenotipi tra individui con uguali genotipi e fattori ambientali, come ad esempio le impronte digitali di due gemelli monozigoti).  $h$  è il path coefficient ( $h$ ), che rappresenta il collegamento tra genotipo e fenotipo, e  $h^2$  è la proporzione di varianza dovuta all’ereditarietà, successivamente chiamata “ereditabilità” in genetica delle popolazioni. Wright, quindi, presentò la path analysis per stimare le relazioni causali tra le variabili, basandosi sulla matrice di correlazione delle variabili osservate, ed enfatizzando i path coefficienti. Inventò, inoltre, un metodo grafico per rappresentare le relazioni causali: il path diagram

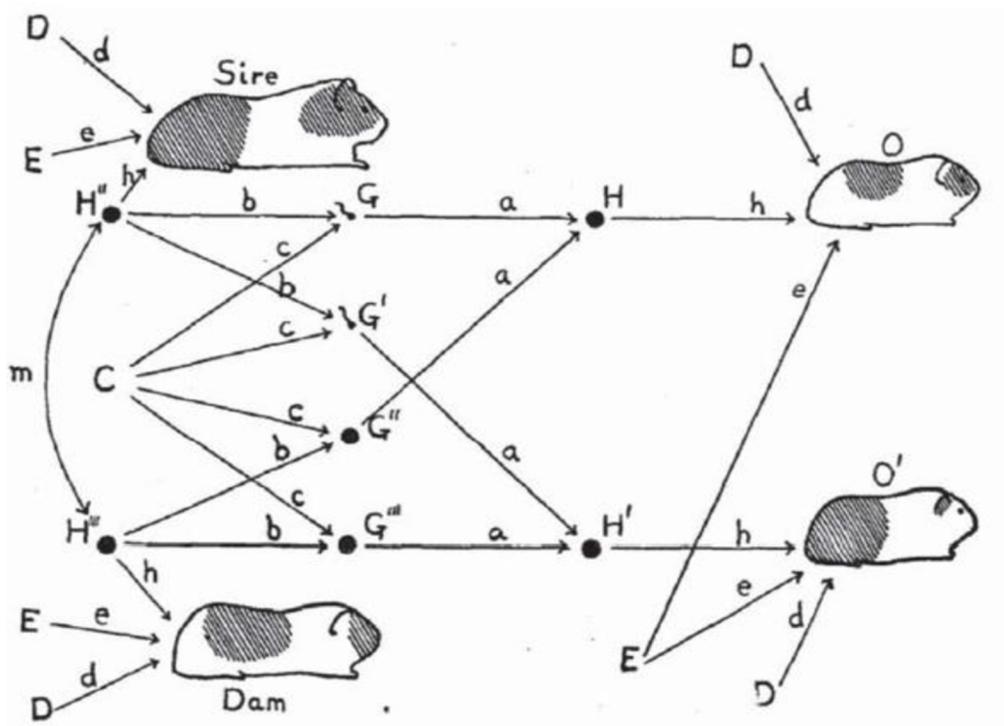


Figura 2. Esempio di Path Diagram di ereditarietà delle cavie, Wright 1921

(Figura2). (Wright 1921)

Il path diagram è costituito da nodi (variabili) connessi da frecce orientate, che rappresentano effetti diretti, o da frecce a due punte che rappresentano correlazioni

non meglio specificate; i *path coefficients* stimati sono riportati sopra le singole frecce.(Wright 1923)

Grazie a questo tipo di rappresentazione grafica è possibile leggere gli effetti totali (diretti e indiretti) e decomporre quindi rapidamente le correlazioni in varie fonti causali.

Nel 1925, lavorando per il dipartimento statunitense dell'agricoltura, sviluppò un complesso sistema ricorsivo di equazioni contenenti variabili osservate e non osservate, temporali e non, per descrivere le possibili correlazioni tra prezzo dei maiali e produzione e prezzo del mais. L'articolo venne però rigettato, e solo alcuni anni dopo pubblicato grazie all'intervento del segretario di stato dell'agricoltura.(2016)

Fino agli anni '50, però, la path analysis fu abbastanza trascurata, finché alcuni econometristi e sociologi non la riconsiderarono per un modello ad 'equazioni simultanee'.

Nel 1972 da Keesling, e nel 1973 da Joreskog e Wiley, venne sviluppato il modello ad equazioni strutturali (SEM) come combinazione tra path analysis e analisi fattoriale confermativa. Conosciuto inizialmente come JKW model e successivamente come LISREL, nome del primo software sviluppato, dal 1994 il modello ad equazione strutturali ebbe un incremento di utilizzo in molte discipline, divenendo negli anni una scelta popolare tra i metodi multivariati, soprattutto in campo sociologico e in econometria.

Questa diffusione nella seconda metà del secolo scorso è dovuta al fatto che questi modelli hanno il vantaggio di poter:

- Sviluppare e testare complicati modelli teorici;
- Considerare nello stesso modello le variabili dipendenti anche come indipendenti;
- Analizzare sia variabili osservate che variabili latenti, ovvero variabili non misurabili direttamente;

## Path Analysis

Sviluppata da Sewall Wright, la Path Analysis è un metodo utilizzato per determinare se un insieme multivariato di dati osservati (non sperimentali) si adatta bene ad un particolare *modello causale*, scelto a priori. (Pedhazur 1998)

È un metodo per scomporre la correlazione osservata tra due variabili, in effetti diretti, indiretti spuri e non analizzabili. (Streiner 2005)

Il Path diagram, che l'output più tradizionale della Path Analysis, permette di visualizzare le relazioni tra le variabili simultaneamente. Potendo, inoltre, distinguere le variabili 'antecedenti', cioè quelle che sono in prima fila nel percorso di casualizzazione che stiamo studiando. Esempi di variabili 'antecedenti' sono l'età e il sesso: tali variabili saranno interpretate come causa di qualcosa ma non saranno interpretate come effetti di altre variabili.

È possibile, inoltre, identificare le variabili '*intercorrenti*', cioè che si inseriscono in posizione intermedia nel percorso, come ad esempio l'ipertensione. Essa può essere causata da variabili come età, sesso e può a sua volta essere causa degli eventi cardiovascolari, ad esempio.

Infine, si possono visualizzare le variabili 'susseguenti', cioè successive a tutte le altre; gli eventi cardiovascolari, potrebbero essere considerati variabili 'susseguenti'.

La Path Analysis è simile ai modelli di regressione tradizionali per i seguenti punti:

- Sia la regressione che la Path Analysis si basano su modelli statistici *lineari*.
- Entrambe le metodologie, presuppongono che le variabili che entreranno far parte del modello siano distribuite in modo gaussiano.

La Path Analysis differisce dai modelli di regressione per i seguenti motivi:

- La Path Analysis è una tecnica multivariata che specifica le relazioni tra le variabili osservate (misurate), risolvendo *simultaneamente* molteplici equazioni per determinare le stime dei parametri. Le variabili nella Path Analysis possono essere interpretate contemporaneamente come indipendenti o dipendenti, mentre le variabili dell'analisi di regressione sono, in ogni equazione, o indipendenti o dipendenti.

- La Path analysis permette inoltre ai ricercatori di riconoscere la imperfezione inevitabile delle loro misure, specificando la varianza inspiegabile; l'analisi di regressione, invece, considerano gli errori di misura come un 'disturbo' ma non permettono di analizzarli in modo esplicito.

## SEM

I Modelli a Equazioni Strutturali (SEM, Structural Equation Models) sono una classe di modelli statistici che consentono di formalizzare le relazioni esistenti tra le variabili, basandosi su una determinata teoria che ipotizza una rete di cause ed effetti.

I SEM, ad esempio, permettono di valutare sia gli effetti diretti ( $A \rightarrow B$ ) che gli effetti indiretti ( $A \rightarrow B \rightarrow C$ ).

Nel caso di effetto indiretto esiste una relazione tra due variabili (A e C), la quale è mediata da una terza variabile (B), che è sia una variabile dipendente (di A), sia indipendente (di C). Abbiamo già sottolineato che il fatto che una variabile possa essere interpretata contemporaneamente sia come variabile dipendente che come variabile indipendente, è un punto importante che distingue la Path Analysis dalla regressione.

Con l'utilizzo dei SEM, si devono introdurre due nuove definizioni di variabile: variabile *esogena* e variabile *endogena*. Esogena è una variabile che svolge sempre e soltanto la funzione di variabile indipendente, mentre la variabile endogena può svolgere sia la funzione di variabile dipendente che indipendente.

I SEM vengono implementati in diversi passi:

- 1- Specificazione del modello;
- 2- Stima dei *parametri strutturali* del modello, ovvero delle relazioni tra le variabili osservate;
- 3- Valutazione del modello, in base all'adattamento ai dati osservati;
- 4- Eventuale modifica del modello.

## 1. SPECIFICAZIONE DEL MODELLO

Nella fase di specificazione del modello occorre individuare le variabili osservate ed eventuali variabili latenti, identificare le variabili esogene ed endogene, e le possibili relazioni tra tutte le variabili prese in considerazione.

Occorre far riferimento ad un modello statistico in grado di formalizzare le relazioni tra le variabili e di definirne i parametri. Il modello di Bentler-Weeks, ad esempio, i parametri da stimare sono i coefficienti di regressione, le varianze e le covarianze delle variabile indipendenti. (Bollen 1989)

## 2. STIMA DEI PARAMETRI

Per la stima dei parametri incogniti possono essere applicati metodi di stima differenti:

- minimi quadrati ordinari (OLS),
- minimi quadrati generalizzati (GLS)
- massima verosimiglianza (ML).

Questi tre metodi possono essere applicati solo quando i dati osservati sono variabili numeriche, discrete e continue, sotto l'assunzione di normalità. La massima verosimiglianza robusta (ML-R), l'Asymptotic Distribution Free (ADF) e il metodo dei minimi quadrati ponderati robusti (WLS-MV) sono alcuni dei metodi sviluppati per la stima in caso di non normalità dei dati e/o di variabili categoriche.

### 3. ADATTAMENTO DEL MODELLO AI DATI

Esistono due differenti strategie per valutare se le stime ottenute si adattano bene ai dati:

- test globali di adattamento all'intero modello
- esaminare l'adattamento dei singoli parametri al modello.

Vi sono molteplici indici di adattamento, che si basano principalmente sulla comparazione della matrice delle covarianze  $\Sigma$  (riprodotta tramite le stime dei parametri) con la matrice di correlazione  $S$  (che stima l'intensità delle relazioni tra le variabili osservate). Se questa differenza è sufficientemente bassa, allora è possibile affermare che i dati si adattano al modello teorico.

Possibili indici di adattamento sono il Root-Mean-Square-Error of Approximation (RMSEA), che considera accettabili valori minori o uguali a 0.05, o il  $\chi^2$ , il quale paragona il modello stimato con un modello saturo, cioè perfetto.

In alternativa è possibile valutare l'adattamento dei singoli parametri, utilizzando tre differenti approcci:

1. Si considera il rapporto tra il parametro stimato e il suo errore standard come valore critico (si assume che il parametro si distribuisca normalmente); se tale rapporto eccede il valore atteso ad uno specifico livello  $\alpha$  (in genere 0.05, test a due code), allora il parametro è significativamente diverso da zero.
2. Si valuta se il segno del parametro coincide con quello atteso dal modello teorico.
3. Si valuta se la stima del parametro è compresa in un range atteso di valori (ad esempio la varianza non deve essere negativa o la correlazione maggiore di 1).

#### 4. MODIFICA DEL MODELLO

Una volta effettuate le verifiche, se l'adattamento del modello non è accettabile bisogna procedere con la modifica e la rivalutazione dei parametri, utilizzando alcune procedure per investigare i specifici errori.

Inizialmente si valuta se i parametri inclusi nel modello hanno un significato pratico, oltre che una significatività statistica. Se è stato stimato significativo un parametro che non si riesce a spiegare, bisogna allora escluderlo dal modello.

Quando si escludono i parametri che non risultano significativi nel modello, bisogna sempre considerare che la significatività dipende dalla potenza del test e, quindi, dalla dimensione campionaria. Se i suddetti parametri sono rilevanti nel modello, essi andrebbero comunque considerati, indipendentemente dalla loro significatività.

Un'alternativa è quella di valutare la matrice dei residui tra osservati e attesi, tenendo presente che valori maggiori di 1.96 o 2.58 indicano che la struttura della covarianza non è ben spiegata dal modello;

Nel 1975, Sörbom propose un metodo basato sull'esame delle derivate prime parziali per i parametri liberi e per quelli fissi. (Sörbom 1975)

## Rischio Cardiovascolare, Aterosclerosi ed infiammazione

Le malattie cardiovascolari (CVD) sono un fenomeno globale e tra le principali cause di morbilità, invalidità e mortalità nel mondo. In Europa rappresentano circa 4 milioni di decessi l'anno, pari alla metà di tutte le cause di morte. Le CVD sono considerate la manifestazione clinica dell'aterosclerosi, processo morboso a carattere progressivo che inizia in genere in età precoce e si manifesta clinicamente durante la mezza e la terza età. (Glass and Witztum 2001; Hollander, Bots et al. 2002; Rosamond, Flegal et al. 2008)

L'aterosclerosi è patologia multifattoriale che, per divenire clinicamente manifesta, richiede la formazione, all'interno della parete di un'arteria, di una placca fibro-lipidica che riduce il flusso ematico. La formazione delle lesioni aterosclerotiche avanzate è la conseguenza di tre processi: 1) l'accumulo di lipidi, principalmente colesterolo libero ed esteri del colesterolo, nello spazio sub-endoteliale delle arterie; 2) l'instaurarsi di uno stato infiammatorio con infiltrazione di linfociti e di macrofagi che, fagocitando i lipidi accumulati, diventano cellule schiumose (foam cell); 3) migrazione e proliferazione di cellule muscolari lisce con produzione di matrice extracellulare.

È importante sottolineare come nelle lesioni vi sia una grande variabilità nel tessuto formatosi a causa della prevalenza di ciascuno dei tre processi. (Stary, Chandler et al. 1994; Stary, Chandler et al. 1995)

In passato, si prendeva in considerazione solo il primo processo, infatti si pensava che l'aterosclerosi fosse una malattia causata solo dal deposito dei lipidi nella parete arteriosa. Oggi, ci sono numerose evidenze cliniche e sperimentali che dimostrano che anche l'infiammazione gioca un ruolo determinante nell'iniziazione, nella progressione e nella destabilizzazione delle placche ateromasiche della parete arteriosa, i cui maggiori costituenti sono tipiche cellule infiammatorie quali cellule mononucleate, macrofagi e linfociti T. (Ross 1999; Pearson, Mensah et al. 2003; Ridker 2003)

Nell'ultima decade è stato evidenziato il ruolo fondamentale dell'infiammazione nella patogenesi dell'aterosclerosi; infatti oggi è definita come una patologia a base infiammatoria. L'infiammazione è il processo che collega gli effetti dell'iperlipemia allo sviluppo della placca, sia nelle fasi iniziali che quelle tardive. Inoltre, è fondamentale nella transizione di placca stabile a placca instabile e quindi nella conversione dell'aterosclerosi da patologia cronica a patologia acuta.

Tuttavia, come per molte innovazioni nella scienza e nella medicina, le radici di questo concetto, apparentemente moderno, risalgono a molto tempo fa. I papiri egiziani di quasi 5.000 anni fa si riferiscono al calore e agli arrossamenti come concomitanti della malattia. Nel I secolo, Aulo Cornelio Celso definì i segni cardinali dell'infiammazione: rossore, gonfiore, calore e dolore. L'avvento del microscopio e dei coloranti all'anilina ha posto le basi per l'era della patologia cellulare e della biologia cellulare dell'infiammazione. Nel XIX secolo, appassionati osservatori hanno descritto la diapedesi dei leucociti dal sangue ai tessuti. Già nel XIX secolo, Rudolf Virchow riconobbe la natura infiammatoria delle placche aterosclerotiche. "In alcuni casi, particolarmente violenti, l'ammorbidimento [delle placche] si manifesta anche nelle arterie non come conseguenza di un processo realmente grasso, ma come prodotto diretto dell'infiammazione.". Tuttavia, l'applicazione pratica di questi concetti allo studio e alla cura dell'aterosclerosi ha subito un ritardo di quasi un secolo. Una volta riconosciuto il ruolo dell'infiammazione nell'aterogenesi, si sono estesi gli studi basati sull'analisi di biomarcatori di infiammazione dagli esperimenti su animali e su cellule coltivate alle osservazioni in-vivo su campioni umani. Si è quindi osservato che biomarcatori di infiammazione, come la proteina C-reattiva (CRP), aumentano negli individui con infarto miocardico acuto. (Liuzzo, Biasucci et al. 1994)

Le prove accumulate negli ultimi 15 anni dimostrano che moderati aumenti dei biomarcatori di infiammazione (come la CRP) sono predittori indipendenti di eventi cardiovascolari prospettici in soggetti apparentemente sani.(Ridker, Cushman et al. 1997)

Alcune linee guida e raccomandazioni delle associazioni professionali incorporano ora l'hsCRP (proteina-C reattiva ad alta sensibilità) in algoritmi di predizione del rischio. Il Risk Score di Reynolds, in particolare, incorpora nel suo algoritmo l'hsCRP e la storia familiare di malattia coronarica prematura, oltre ai tradizionali fattori di rischio. (Ridker, Buring et al. 2007)

Nonostante che l'attenzione si sia focalizzata sulla Proteina C-Reattiva (PCR), a causa delle sue caratteristiche biologiche di stabilità, della relativa facilità di determinazione e dell'esistenza in commercio di metodi analitici standardizzati ad alta sensibilità e di costo moderato, sono numerosi i mediatori dell'infiammazione che giocano un ruolo fondamentale nell'aterosclerosi.

Tra di essi annoveriamo:

- **Interleuchina-6** → L'Interleuchina-6 (abbreviato IL-6) è una citochina pleiotropica, prodotta da diversi tipi di linfociti, che svolge un ruolo importante in una vasta gamma di risposte quali le reazioni immunitarie, le reazioni di fase acuta ed l'ematopoiesi. IL-6 agisce come una citochina multifunzionale, sia pro-infiammatoria, sia antinfiammatoria. È prodotta da una varietà di tipi di cellule, le fonti principali sono i macrofagi, i fibroblasti e le cellule endoteliali. (Naka, Nishimoto et al. 2002). IL-6 è un polipeptide di 26 kDa, organizzato in domini globulari di 4  $\alpha$ -eliche a lunga catena, in cui la prima e la seconda sono parallele fra loro ed antiparallele alla terza e alla quarta (Heinrich, Behrmann et al. 1998). Questa molecola attiva le vie intracellulari tramite un complesso recettoriale che comprende due unità trasducenti  $\beta$ , dette gp130, ed una subunità specifica  $\alpha$ , IL-6R, che è l'unica con cui la citochina può interagire direttamente e che è fondamentale per la trasduzione del segnale.

L'Interleuchina-6 stimola l'attivazione endoteliale, la proliferazione delle cellule muscolari lisce ed il reclutamento dei leucociti. È inoltre coinvolta nell'alterato metabolismo lipidico e nella produzione dei trigliceridi, rappresentando un importante fattore di rischio per le malattie cardiovascolari. IL-6 riduce l'attività

della lipoproteina lipasi ed i livelli monomerici di questa nel plasma, mentre aumenta la captazione lipidica dei macrofagi.

I macrofagi attivati esprimono IL-6, suggerendo un ruolo importante di questa citochina nella progressione del processo aterosclerotico. Durante il danno vascolare le cellule muscolari lisce sono esposte alle piastrine ed ai loro prodotti e la produzione di citochine da parte delle cellule muscolari lisce probabilmente contribuisce al danno vascolare. Inoltre l'IL-6 circolante stimola l'asse ipotalamo-ipofisi-surrene la cui attivazione è associata con obesità centrale, ipertensione ed insulino-resistenza. Per questi motivi è stato proposto un ruolo per la IL-6 nella patogenesi delle malattie cardiovascolari attraverso una combinazione di meccanismi autocrini, paracrini ed endocrini. Biasucci et al. suggeriscono che gli elevati livelli di IL-6 sono predittivi di futuri eventi cardiovascolari e di un esito infausto negli eventi coronarici acuti. (Biasucci, Liuzzo et al. 2000)

- **Solubile Glycoprotein 130** (sgp130) → è considerato un inibitore naturale della risposta dell' interleuchina-6, dipendente dalla suo recettore solubile (sIL-6R). Quindi, possiamo considerare sgp130 come un'antagonista del recettore dell'interleuchina-6, ovvero una molecola anti-infiammatoria. Recentemente è stato dimostrato, nei topi, che la somministrazione di sgp130 riduce significativamente lo sviluppo e la progressione della placca aterosclerotica, neutralizzando il complesso IL-6/sIL-6R. (Schuett, Oestreich et al. 2012)
- **Globuli bianchi** → ( o leucociti, in inglese white blood cells, WBC) sono un tipo di cellula del sangue che si forma nel midollo osseo e si trova nel sangue e nei tessuti linfatici. La loro funzione principale è quella di preservare l'organismo dagli attacchi di microrganismi patogeni quali batteri o virus, migrando nel sangue per mezzo di agenti chemio-tattici che permettono loro di raggiungere la sede dell'infiammazione. Svolgono inoltre funzione di difesa contro corpi estranei penetrati nell'organismo previo superamento delle barriere costituite dalla cute e dalle mucose. In condizioni normali la loro concentrazione nel sangue è di circa  $7000/\text{mm}^3$ . Un numero di leucociti inferiori alla norma viene

detta leucopenia, invece un numero eccessivo, e superiore alla norma viene detto leucocitosi. La presenza di un numero superiore alla norma di leucociti, in genere, può essere causata da infezioni o infiammazioni. Se ad aumentare sono i granulociti neutrofili, si tratta di un'infezione di natura batterica, se invece ad aumentare sono i linfociti, allora l'origine è di tipo virale; se invece aumentano i granulociti eosinofili, in genere si tratta di un'allergia o di parassiti. I leucociti si distinguono in granulociti e agranulociti, a seconda che ci sia presenza o assenza di granuli nel citoplasma. I granulociti comprendono:

- granulociti neutrofili,
- granulociti eosinofili,
- granulociti basofili.

Un'importante caratteristica funzionale dei leucociti è la motilità. Tali cellule, infatti, non sono solo trasportate passivamente nel torrente circolatorio, ma sono anche in grado di sviluppare una locomozione attiva utilizzando alcune proteine del citoscheletro.

## **Markers dell'aterosclerosi: spessore del complesso medio-intimale (IMT)**

Lo spessore del complesso medio-intimale carotideo o C-IMT, (dall'inglese, Carotid Intima-media Thickness), è oggi considerata uno dei markers di patologia aterosclerotica subclinica più precisi ed accettati dalla comunità scientifica. Sin da metà dagli anni '80, l'ultrasonografia B-Mode stata utilizzata per valutare la parete dei vasi carotidei e le eventuali placche presenti.

La versatilità di questa metodica è risultata essere particolarmente utile per la stratificare meglio i soggetti asintomatici che, secondo gli attuali algoritmi per il calcolo del rischio cardiovascolare globale, risultavano essere a rischio intermedio (Bernetti, Abbate et al. 2011). La metodica permette infatti di evidenziare e quantificare la patologia aterosclerotica sin dalle prime fasi dove è possibile osservare piccole alterazioni medio-intimali della parete vascolare.

Nel 1986 Pignoli et al. a dimostrarono che la classica doppia riga che si vede a livello di una parete arteriosa visualizzata con ultrasonografia B-mode corrisponde al complesso medio intimale.(Pignoli, Tremoli et al. 1986). In studi successivi fu possibile rilevare che variazioni del C-IMT si associava ai fattori di rischio dell'aterosclerosi.

In una arteria sana, il C-IMT è costituito per l'80% da tonaca media e per circa un 20% da tonaca intima, comprendente endotelio, lamina basale e matrice subendoteliale (*Figura 3*).

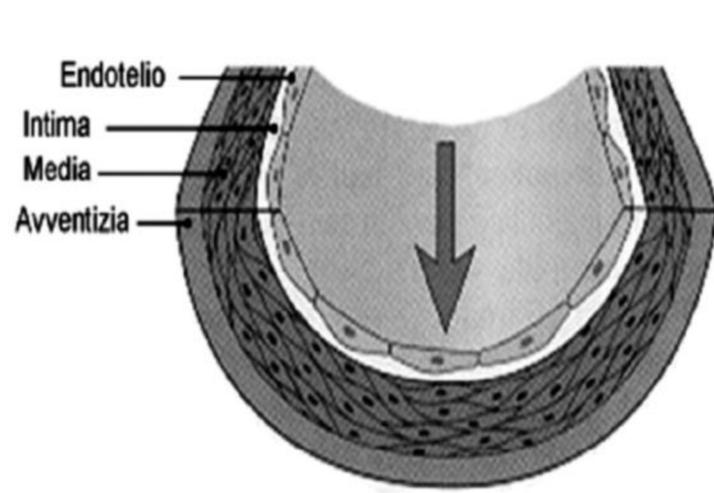


Figura 3. Anatomia arteria Carotide

I valori di C-IMT in zone libere da placca variano da 0.25 a 1.5 mm negli adulti normali. Studi hanno dimostrato che, in individui sani il C-IMT correla con l'età, con un aumento medio di spessore pari a circa 0.03/0.04 mm ogni dieci anni.(Ebrahim, Papacosta et al. 1999)

La carotide è uno dei maggiori tronchi arteriosi del corpo umano e irrorà, insieme all'arteria vertebrale, il sistema nervoso centrale e le strutture facciali.

Dal punto di vista anatomico, il vaso può essere suddiviso in tre porzioni: carotide comune, carotide interna e carotide esterna.

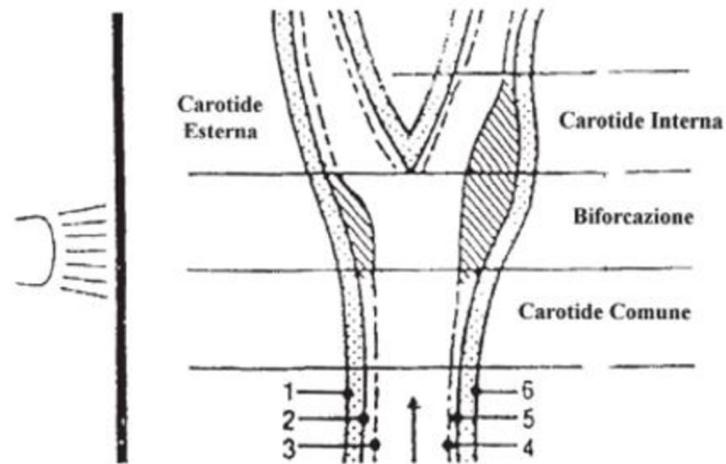


Figura 4. Anatomia arteria Carotide

La vascolarizzazione del sistema nervoso centrale è garantita dal poligono del Willis (Figura 5), un sistema di anastomosi arteriose alla base della scatola cranica, costituito dalla confluenza di arteria basilare (costituita da arteria vertebrale destra e sinistra) e dalle due carotidi interne (destra e sinistra).

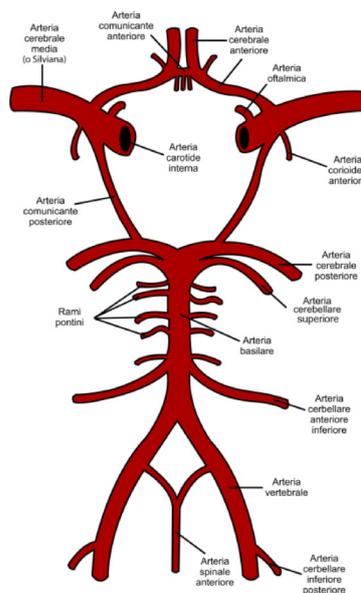


Figura5: Poligono del Willis

Il sistema anastomotico circolare del poligono del Willis è coinvolto nell'irrorazione dei tessuti cerebrali anche in caso di ostruzione di uno dei rami affluenti. In alcuni casi di occlusione vascolare si può verificare anossia di una o più parti del territorio di distribuzione vascolare (ICTUS) a valle. Diventa così evidente l'utilità della valutazione preventiva volta a verificare lo sviluppo della patologia aterosclerotica vascolare a carico dei vasi carotidei ed in particolare per l'individuazione dei pazienti che, pur essendo asintomatici, sono comunque a rischio di sviluppare la patologia aterosclerotica; condizione predisponente per gli eventi cardio e cerebrovascolari.

La misurazione del C-IMT può essere effettuata in differenti aree dell'albero carotideo destro e sinistro in una o più porzioni del vaso (carotide comune, bulbo o carotide interna) valutando sia la parete più profonda (in inglese: far wall) che la parete più superficiale (near wall) rispetto al piano cutaneo. Nelle analisi statistiche si usa in genere la media dei massimi valori rilevati in tutti i segmenti analizzati ( $C-IMT_{mean-max}$ ), o il valore massimo rilevato a livello della placca

Alcuni studi hanno dimostrato come ci sia una relazione tra la progressione del C-IMT (variazione nel tempo) e il rischio di eventi coronarici (RR aumenta di circa il 15% per ogni aumento di 0,10 mm dello C-IMT carotideo) e cerebrovascolari. (Baldassarre, Veglia et al. 2013). La presenza di una placca aterosclerotica, definibile come un valore di IMT >1.5 mm o un aumento focale del C-IMT >50% rispetto al valore di C-IMT rilevato nei segmenti limitrofi liberi da placca, aumenta il rischio di arteriopatia coronarica quale che sia il livello di C-IMT in tutti gli altri distretti carotidei considerati.

I valori di C-IMT carotideo >1 mm, sono stati correlati ad un aumento di 2 volte del rischio di infarto miocardico acuto a 3 anni. (Simon, Garipey et al. 2002)

Valori aumentati di C-IMT sono indice attendibile dello stato aterosclerotico di altri distretti vascolari. Valori superiori a 0.8 mm si associano, infatti, ad una probabilità di presenza di placche aterosclerotiche in altre aree dell'albero carotideo (ed in altri distretti vascolari) pari al 95.5%, mentre con uno spessore tra 0.6 - 0.8 tale probabilità scende al 53% e si riduce ulteriormente fino al 4.5% quando il C-IMT < 0.6 mm.

La valutazione del C-IMT carotideo consente di ottenere una vera e propria istantanea dell'effetto di tutti i fattori che influenzano la patologia cardiovascolare, sia quelli di rischio che quelli protettivi. Il C-IMT è indipendentemente associato alla maggior parte dei fattori di rischio CV convenzionali quali l'ipertensione arteriosa, l'ipercolesterolemia ed il diabete; ambientali come il fumo di sigaretta ed emergenti quali l'insulino-resistenza, e l'iperuricemia; e ultimo ma non ultimo con fattori protettivi come i livelli di colesterolo HDL, la longevità dei genitori ecc.

Uno dei limiti associati a questa metodica è la variabilità che si riscontra nella precisione e l'accuratezza della misura fra diversi laboratori. Numerosi sono i fattori che possono influenzare la variabilità della misura del C-IMT, fra questi il tipo di apparecchiatura utilizzata per l'acquisizione e/o la lettura delle immagini, l'anatomia del paziente (profondità e presenza di particolari tortuosità) e all'esperienza degli operatori (sonografisti e lettori).

Il training ed il controllo degli operatori sono determinanti importanti per ottenere misure accurate e riproducibili.

## Eventi vascolari maggiori

### **Malattia cardiovascolare (CVD)**

La CVD, la principale complicanza dell'aterotrombosi, è la prima causa di morte a livello mondiale, con un aumento annuale dei decessi da 17,3 a 23,6 milioni all'anno entro il 2030. Attualmente i decessi per CV rappresentano circa 1/3 di tutti i decessi a livello mondiale. Le spese mediche e i costi dei farmaci associati al trattamento del CVD sono enormi, per un totale di 126 miliardi di dollari nel 2010, nei soli Stati Uniti. (Odden, Coxson et al. 2011). L'attenzione dedicata alla CVD è dovuta al rapido aumento dei decessi dovuti a questa patologia negli ultimi decenni (incremento del 12,5%). In realtà, nel decennio 2005/2015 si è registrata una riduzione del 15,6% dei decessi specifici per età, anche se l'aumento della speranza di vita ha prodotto un incremento assoluto dei decessi per CVD. (Norheim, Jha et al. 2015; Roth, Forouzanfar et al. 2015; Roth, Johnson et al. 2017) Inoltre, ciò è stato rilevante nei Paesi ad alto reddito, rispetto ai Paesi a medio reddito e ai Paesi a basso reddito. (1999; Roth, Forouzanfar et al. 2015) Di conseguenza, la maggioranza dei decessi da CVD avviene ora nei Paesi a medio e basso reddito.

Nel 2013, l'Organizzazione Mondiale della Sanità ha lanciato il "25x25 Global Action Plan", un progetto volto a ridurre la mortalità del 25% entro il 2025. Questo progetto si concentra sulla correzione di 4 principali comportamenti legati alla salute: consumo di tabacco, dieta, attività fisica e alcol. L'insorgenza di CVD è multifattoriale e tra i principali fattori di rischio vi sono il diabete mellito di tipo 2 (T2DM), l'ipertensione, l'abitudine al fumo, il sovrappeso e la dislipidemia. Il riconoscimento precoce e il trattamento dei pazienti ad alto rischio di aterosclerosi è uno degli obiettivi principali per ridurre l'incidenza di eventi aterotrombotici e malattie cardiovascolari.

### **Malattia cerebrovascolare**

L'ictus è il secondo tipo di evento che contribuisce maggiormente al carico di malattie vascolari. La prevalenza dell'ictus aumenta con l'età, con un picco tra i soggetti con età compresa tra i 74 e i 79 anni. Anche la mortalità legata all'ictus aumenta con l'età. Circa il 63% dell'ictus ischemico e l'80% dell'ictus emorragico si verifica ora nei Paesi a

medio e basso reddito. L'ictus ischemico è più comune dell'ictus emorragico, sebbene l'ictus emorragico sia associato a una maggiore mortalità e contribuisca maggiormente alla disabilità. In media, le popolazioni dei Paesi a medio e basso reddito soffrono di ictus ad un'età di 6 anni minore rispetto ai Paesi ad alto reddito. (Feigin, Forouzanfar et al. 2014) I paesi con la più alta prevalenza di ictus, standardizzando per età, sono attualmente in Oceania, Europa orientale, Asia centrale e Sud-est asiatico, mentre la mortalità da ictus è più elevata in Oceania e Africa subsahariana. (Feigin, Forouzanfar et al. 2014)

**TIA.** In un'epoca caratterizzata da un'ampia diffusione della risonanza magnetica per immagini, la definizione storica di TIA come "un episodio transitorio di disfunzione neurologica causata da una disfunzione cerebrale focale, midollo spinale o ischemia retinica senza infarto acuto" (Easton, Saver et al. 2009) che dura meno di 24 ore, non è più appropriata. Sebbene la maggior parte dei pazienti siano asintomatici alla presentazione e abbiano un normale esame fisico, (Edlow 2018) l'incidenza dell'ictus ischemico acuto entro 48 ore dalla visita di emergenza per la TIA è del 4,8% (182/3.814 pazienti; 95% CI dal 4,0% al 5,6%).(Giles and Rothwell 2007) La prevenzione dell'ictus secondario - la rapida implementazione di interventi multipli - riduce l'esito dell'ictus dell'80%. (Rothwell, Giles et al. 2007; Stead, Bellolio et al. 2009; Amarenco, Lavalley et al. 2016) Poiché una diagnosi accurata, test rapidi e implementazione dei trattamenti possono prevenire la disabilità e gli ictus fatali, la TIA ha un'enorme importanza per la salute pubblica. (Edlow 2018)

## **Malattie vascolari periferiche**

Le PAD colpiscono più del 5% della popolazione di età superiore ai 60 anni. Negli Stati Uniti, le PAD colpiscono circa 8-10 milioni di americani e causano ogni anno 500.000 ricoveri ospedalieri e l'esecuzione di 100.000 angiogrammi. (Davies 2012) Una malattia coronarica (CAD) significativa (in almeno un'arteria coronarica) è stata documentata nel 60% - 80% dei pazienti con PAD, e una stenosi carotidea emodinamicamente significativa (mediante ecografia Doppler) è stata riscontrata nel 12% - 25% dei casi.(Lambert and Belch 2013) Il rischio di AMI è aumentato del 20% - 60%, mentre il rischio di morte coronarica è aumentato da 2 a 6 volte nei pazienti con PAD. Per questo motivo, la diagnosi precoce di PAD può aiutare a identificare i pazienti a rischio

di eventi CV. (Hirsch, Haskal et al. 2006) La PAD è associata ad un aumento del 40% del rischio di ictus e la gravità della PAD è associata positivamente all'incidenza di TIA e ictus. (Davies 2012) Lo spettro clinico della PAD è molto variabile: i pazienti possono essere asintomatici o provare dolore anche in seguito ad una deambulazione minima. La maggior parte dei pazienti asintomatici con PAD sarà identificata attraverso lo screening dell'indice ankle-brachial index (ABI). (Leng, Fowkes et al. 1996) L'ABI (cioè il rapporto tra la pressione arteriosa sistolica dell'arto inferiore e quella del braccio) è un mezzo semplice e affidabile per valutare la presenza e gravità della PAD. (Norgren, Hiatt et al. 2007) La PAD con *claudicatio intermittens* (IC) spesso non è diagnosticata e, a sua volta, sottotrattata. La bassa percentuale di diagnosi (~ 30%) tra le persone con PAD è preoccupante, a causa dell'alto rischio di esiti avversi legati a un eventuale peggioramento della PAD. (Newman, Shemanski et al. 1999; Resnick, Lindsay et al. 2004; van der Meer, Bots et al. 2004) Tra i soggetti con ABI <0.9, ~ Il 25% sperimenta un peggioramento della claudicazione che richiede un intervento chirurgico o un'amputazione. In un follow-up a 10 anni, circa il 55% dei pazienti PAD (ABI <0.9) è morto di malattia CV, il 10% di malattia cerebrovascolare, e il 25% di ragioni non vascolari. Meno del 10% è morto per altri eventi vascolari periferici (soprattutto aneurismi aortici nell'addome). In un altro follow-up,(Criqui 2001) la mortalità a 10 anni era del 61,8% nei maschi con PAD sintomatico (ABI <0,9); in confronto, nei maschi senza PAD (ABI = 0,9) la mortalità era del 16,9%. I tassi di mortalità nelle femmine sono stati rispettivamente del 33,3% e dell'11,6%. Meno del 25% dei pazienti con PAD (ABI < 0.9) è sopravvissuto per 10 anni, essendo la mortalità vascolare la causa dominante di morte in tale contesto. Dopo la correzione per i classici fattori di rischio, la gravità della PAD era un fattore predittivo indipendente della mortalità.

L'ischemia critica degli arti (CLI; ABI<0.4) è osservata nel 12% della popolazione di PAD. Nella maggior parte dei casi, la CLI è una complicanza trombotica avanzata della PAD, dovuta ad un inadeguato flusso di sangue a riposo agli arti inferiori, ed è caratterizzata da dolore a riposo, ulcerazione, ed eventualmente cancrena e perdita dell'arto.(O'Connell and Quiñones-Baldrich 2009). La CLI è raramente il risultato di un evento acuto (ad es. embolia, trombosi o trauma). I pazienti con CLI sono candidati ad una rapida rivascolarizzazione. La CLI è associata ad un aumento di mortalità: nel primo anno dopo la diagnosi, il 25% dei pazienti (il 45%, se hanno già subito

un'amputazione) morirà, e il 30% di loro avrà amputazioni, mentre solo il 45% sopravvivrà con entrambe le gambe. Dopo 5 anni oltre il 60% dei pazienti è morto.

## Razionale

Ad oggi, nella ricerca biomedica, la metodologia statistica implementata per studiare l'associazione tra un marcatore e uno stato di malattia è l'analisi della regressione. Questo tipo di approccio coinvolge una singola variabile dipendente e più variabili indipendenti, non rilevando relazioni complesse tra le variabili.

La Path Analysis ed i modelli ad equazione strutturale (SEM) sono metodologie statistiche innovative che permettono di studiare e visualizzare le relazioni più significative tra l'intero set di variabili selezionate, valutando simultaneamente più relazioni.

Queste metodologie possono essere utilizzate in ambito esplorativo, identificando simultaneamente l'intero insieme delle relazioni, dirette e indirette, esistenti tra le variabili. È possibile visualizzare queste relazioni tramite un "path diagram".

Oppure, possono essere utilizzate in ambito confermativo, in quanto capaci di fornire stime ed indici di bontà di adattamento dei modelli ipotizzati.

Attraverso l'utilizzo di questi modelli statistici avanzati, si vuole focalizzare l'attenzione sulle relazioni che intercorrono tra i biomarcatori dello stato infiammatorio e l'aterosclerosi subclinica, valutata mediante differenti misurazioni di spessore medio-intimale carotideo (IMT), ed eventi vascolari.

Path Analysis e SEM permettono di comprendere meglio le relazioni tra i differenti fattori di rischio analizzati, i biomarcatori di stato d'infiammazione e c-IMT, potendo valutare sia gli effetti diretti che quelli indiretti tra le variabili.

L'attenzione è stata focalizzata sui biomarcatori di stato di infiammazione dato il ruolo centrale dell'infiammazione sistemica di basso grado sia nell'eziopatogenesi dell'aterosclerosi subclinica che nell'insorgenza degli eventi cardiovascolari. Diversi studi hanno analizzato l'associazione tra infiammazione ed eventi cardiovascolari e/o marcatori di aterosclerosi subclinica, caratterizzando l'infiammazione sistemica come alterazione dei valori della proteina C-reattiva (CRP), dei globuli bianchi (WBC), oppure delle interleuchine infiammatorie, come ad esempio l'interleuchina-6. Altri studi

utilizzano l'abitudine al fumo e il numero di sigarette fumate come un indice di infiammazione. Raramente, tuttavia, sono stati studiati gli effetti di molteplici biomarcatori di infiammazione, considerati congiuntamente, sugli indici di aterosclerosi subclinica e sugli eventi cardiovascolari. Per le analisi verrà utilizzato il database dello studio IMPROVE, che comprende 3703 soggetti ad alto rischio cardiovascolare, arruolati in 7 centri differenti di 5 città europee.

## **Obiettivo dello studio**

Gli obiettivi principali del presente studio erano:

- Confrontare metodologie statistiche innovative, quali la Path Analysis e i SEM, con metodologie statistiche 'classiche' in uno studio multicentrico, longitudinale, osservazionale. In particolare, valutare come tali metodologie siano in grado di caratterizzare la rete delle relazioni tra fattori di rischio, classici ed emergenti, e l'aterosclerosi subclinica, valutata mediante il c-IMT.
- Rilevare la posizione intermedia del c-IMT (e della sua variazione nel tempo) nella catena causale che conduce dai fattori di rischio agli eventi cardio- e cerebro-vascolari
- Valutare, con queste metodologie innovative, la relazione tra uno score che rifletta lo 'stato' infiammatorio dei pazienti e c-IMT, la sua progressione e gli eventi vascolari maggiori.

# **Metodi**

## **Popolazione**

### **Lo studio IMPROVE**

I dati utilizzati nella presente tesi si riferiscono ai pazienti reclutati nello studio IMPROVE, uno studio multicentrico, longitudinale ed osservazionale che coinvolge 7 centri di 5 paesi europei: Finlandia, Svezia, Olanda, Francia, Italia.

L'obiettivo primario dello studio IMPROVE era valutare se la variazione del C-IMT nel tempo, misurata nei primi quindici mesi di studio, fosse un predittore indipendente di rischio di eventi vascolari (AMI, morte cardiovascolare, ictus o qualsiasi intervento di rivascolarizzazione nei distretti carotideo, coronarico o periferico) verificatisi nei successivi 21 mesi (dal 15° al 36° mese) di follow-up.

Lo studio prevedeva l'arruolamento di 3.598 pazienti (514 per centro). A tal fine sono stati esaminati circa 21.000 soggetti: 3.400 a Milano, 1.450 nel primo centro Kuopio e 2.354 nel secondo, 4.239 a Stoccolma, 4.050 in Olanda, 3.804 a Perugia e 1.800 a Parigi. Erano considerati idonei soggetti di entrambi i sessi, con età compresa tra 55 e 79 anni, asintomatici per malattie cardiovascolari e non portatori di patologie che potevano limitare la longevità del paziente e da situazioni che potevano limitare una corretta visualizzazione c-IMT (ad esempio eccessiva obesità).

Un importante fattore di inclusione era la presenza di almeno tre fattori di rischio (VRF).

I soggetti erano considerati esposti ad un VRF quando era soddisfatto uno dei seguenti criteri:

- sesso maschile o almeno 5 anni dopo la menopausa per le donne;
- ipercolesterolemia (livelli medi di LDL-C calcolati  $> 160$  mg/DL o trattamento con farmaci per l'abbassamento dei lipidi);
- ipertrigliceridemia (livelli di trigliceridi  $> 200$  mg/DL dopo dieta o trattamento con farmaci per l'abbassamento dei trigliceridi);

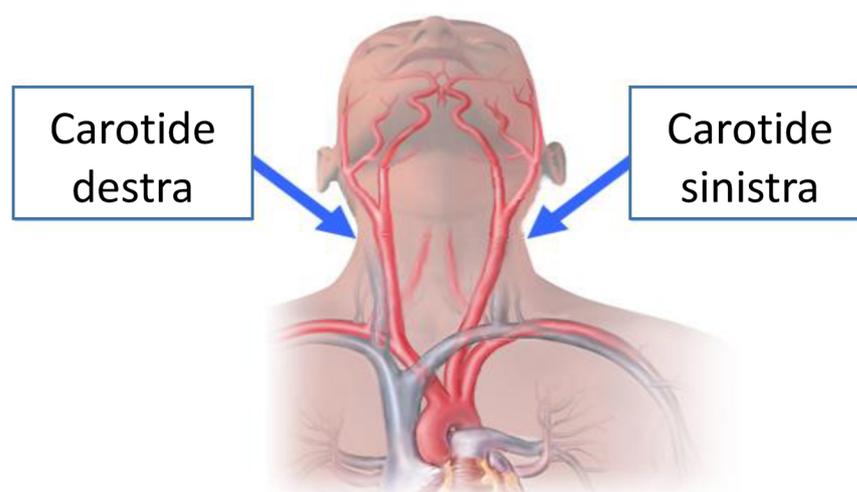
- ipoalfalipoproteinemia (HDL-C < 40 mg/DL);
- ipertensione (pressione sanguigna diastolica > 90 mmHg e/o pressione sanguigna sistolica >140 mmHg o trattamento con farmaci antiipertensivi);
- diabete o glicemia a digiuno compromessa (livello di glicemia > 110 mg/DL o trattamento con insulina o farmaci ipoglicemici orali);
- abitudini di fumo (almeno 10 sigarette/giorno per almeno trenta mesi);
- anamnesi familiare di malattie cardiovascolari.

I criteri di esclusione erano: età <54 anni o >79 anni; configurazione anatomica anormale dei muscoli del collo, marcata tortuosità e/o eccessiva profondità dei vasi carotidei; storia personale di AMI, angina pectoris, ictus, TIA, aneurisma aortico, claudicazione intermittente, intervento chirurgico di rivascolarizzazione nei distretti arteriosi carotidei, coronarici o periferici, scompenso cardiaco (III-IV classe NYHA); e storia di condizioni mediche gravi che potessero limitare la longevità (ad esempio neoplasie).

In totale sono state reclutate 3.711 persone. Tutti i partecipanti hanno firmato due consensi informati (uno per la partecipazione generale allo studio e uno per gli studi di genetica).

## Misura dello spessore del complesso medio-intimale (IMT) e sua variazione nel tempo

L'ecografia carotidea è stata eseguita utilizzando sette ecografi identici (sistema Technos, ESAOTE), dotati di sonde lineari da 5-10 MHz. Ecografisti e lettori sono stati formati e certificati dal Centro di coordinamento (Dipartimento di Scienze Farmacologiche e Biomolecolari) di Milano. In tutti i pazienti sono state analizzate: la carotide comune destra e sinistra (CC), la biforcazione (BIF) e l'arteria carotide interna (ICA) in tre diverse proiezioni (anteriori, laterali e posteriori) e registrate su videocassette sVHS.



Per le analisi statistiche sono state considerate undici variabili di IMT carotideo ed in particolare: gli spessori medi ( $1^{\text{st}}\text{CC-IMT}_{\text{mean}}$ ) e massimi ( $1^{\text{st}}\text{CC-IMT}_{\text{max}}$ ) del primo centimetro della carotide comune misurato in zone libere da placca; della restante parte della carotide comune ( $\text{CC-IMT}_{\text{mean}}$  e  $\text{CC-IMT}_{\text{max}}$ ); della biforcazione ( $\text{BIF-IMT}_{\text{mean}}$  e  $\text{BIF-IMT}_{\text{max}}$ ), e della carotide interna ( $\text{ICA-IMT}_{\text{mean}}$  e  $\text{ICA-IMT}_{\text{max}}$ ). Queste variabili sono state poi convogliate in variabili riassuntive relative all'intero albero carotideo e in particolare il  $\text{C-IMT}_{\text{mean}}$ , (media degli IMT rilevati in tutti i segmenti carotidei considerati),  $\text{C-IMT}_{\text{max}}$  (dimensione del più grosso valore di IMT rilevato nell'intero albero carotideo) e  $\text{C-IMT}_{\text{mean-max}}$  (media dei valori massimi rilevati in tutti i segmenti carotidei considerati). In particolare:

$\text{IMT}_{\text{mean}}$  è la media di  $1^{\text{st}}\text{CC-IMT}_{\text{mean}}$ ,  $\text{CC-IMT}_{\text{mean}}$ ,  $\text{BIF-IMT}_{\text{mean}}$  e  $\text{ICA-IMT}_{\text{mean}}$ .

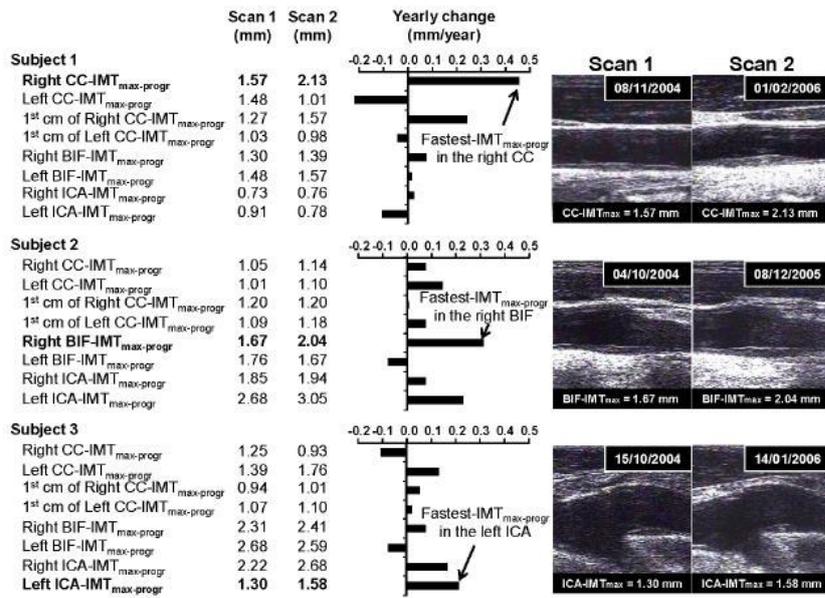
L' $IMT_{max}$  è il valore più elevato tra 1<sup>st</sup>  $CCC-IMT_{max}$ ,  $CC-IMT_{max}$ ,  $BIF-IMT_{max}$  e  $ICA-IMT_{max}$ .

$IMT_{mean-max}$  è la media di 1<sup>st</sup>  $CC-IMT_{max}$ ,  $CC-IMT_{max}$ ,  $BIF-IMT_{max}$  e  $ICA-IMT_{max}$ .

Le analisi effettuate nell'ambito della presente tesi sono state effettuate utilizzando come variabile di interesse lo  $IMT_{mean-max}$ . Tale scelta è basata su evidenze sperimentali che dimostrano che questa variabile è quella meglio associata agli eventi cardio e cerebrovascolari. (Baldassarre, Nyssonen et al. 2010)

Per tutte le variabili di C-IMT sopra descritte è stata valutata anche la variazione avvenuta nell'ambito dei primi 15 mesi di studio. Oltre a queste variabili è stata anche calcolata la cosiddetta "Fastest- $IMT_{max}$ -progression". Per quanto riguarda le analisi riguardanti le variazioni di IMT nel tempo, nella presente tesi si è deciso di utilizzare come variabile di interesse quest'ultima variabile in quanto evidenze sperimentali hanno documentato essere l'unica variabile di variazione di IMT nel tempo ad essere un ulteriore predittore indipendente di eventi vascolari. (Baldassarre, Veglia et al. 2013)

La Fastest- $IMT_{max}$ -progression è stata calcolata come il massimo tra tutte le progressioni rilevante in tutti i segmenti carotidei considerati, ossia il valore massimo fra la progressione del  $CC-IMT_{max}$ , del  $Bif-IMT_{max}$ , dell' $ICA-IMT_{max}$ ) vedi figura 6.



Supplemental Figure IV: Baseline and 15<sup>th</sup>-month C-IMT values of three subjects who showed the Fastest-IMT<sub>max-progr</sub> in: common carotid (Subject 1), bifurcation (Subject 2), and internal carotid artery (Subject 3).

Figura 6: Tratta da supplementary online material di Baldassarre et al. (Baldassarre, Veglia et al. 2013)

## Analisi di laboratorio

Il prelievo di sangue per gli esami di laboratorio è stato effettuato dopo 12 ore di digiuno. La concentrazione plasmatica di glucosio, la conta delle cellule del sangue, l'ematocrito e la formula leucocitaria sono stati misurati in laboratori locali. Il sangue per le analisi biochimiche centralizzate è stato mantenuto a temperatura ambiente per un minimo di 30 minuti per consentire la coagulazione. Il siero, preparato mediante centrifugazione a 2000g per 20 minuti, è stato aliquotato in tubi di polipropilene e conservato a -80°C. I campioni sono stati poi spediti a Stoccolma (Karolinska Institute Stockholm, Svezia) in ghiaccio secco. I campioni biologici sono stati poi utilizzati per le analisi biochimiche centralizzate e per la conservazione nella biobanca.

Le analisi biochimiche centralizzate comprendevano la misurazione delle concentrazioni sieriche di colesterolo totale, HDL (le LDL sono state calcolate secondo la formula di Friedewald), trigliceridi, acido urico, proteina C-reattiva ad alta sensibilità (hs-CRP) e creatinina. Queste variabili ematologiche sono state misurate presso il Dipartimento di Chimica Clinica del Karolinska University Hospital (Solna, Stoccolma, Svezia) utilizzando strumenti LX Beckman. In particolare, colesterolo, trigliceridi e acido urico sono stati misurati con metodi enzimatici. La creatinina con un metro a colorimetria-alcalina *picric* e la hs-CRP con la turbidimetria (metodo NIPIA).

I livelli di IL6 nel siero sono stati misurati utilizzando il Meso Scale Discovery Human Cytokine Assay (Gaithersburg, MD, USA). I campioni di siero sono stati diluiti 1:2 e le concentrazioni, derivate da una curva standard, sono state riportate sia in pg/mL. Il limite inferiore di rilevazione era 0.06 pg/mL.

Il livello sierico di Sgp130 è stato misurato mediante un metodo elisa (Human sgp130 DuoSet ELISA development kit; #DY228) prodotto da R&D Systems® (R&D systems Minneapolis, MN, USA). Al fine di ottimizzare il saggio sono stati effettuati diversi esperimenti preliminari. Le piastre a 96 pozzetti sono state rivestite con anticorpo diluito (4,0 µg/mL) e incubate per una notte a temperatura ambiente. Dopo il blocco della reazione, effettuato aggiungendo sieroalbumina bovina al 2%, sono stati aggiunti gli anticorpi di rilevazione (0,08 µg/mL) e, in uno step successivo, Streptavidina più il

reagente per colorare il substrato. Infine, la piastra è stata letta mediante con un lettore specifico (Tecan, Infinite F200) impostato su una lunghezza d'onda di 450 nm. Le concentrazioni di Sgp130, derivate dalla curva standard, sono state espresse in ng/mL.

## Analisi statistica

Le variabili quantitative sono state rappresentate come media  $\pm$  deviazione standard (SD), o mediana (intervallo interquartile). Le variabili con distribuzioni non normale sono state trasformate in logaritmo prima dell'analisi.

La relazione tra i Biomarcatori dello stato infiammatorio tra di loro è stata valutata con correlazione di Spearman. I dati sono rappresentati come il coefficiente di correlazione  $r$  e il p-value.

La regressione lineare è stata utilizzata per analizzare l'associazione tra il c-IMT ed i biomarcatori di stato d'infiammazione. Sono stati implementati due modelli: Modello 1, non aggiustato; Modello 2, includendo i 4 biomarcatori dello stato infiammatorio nello stesso modello multivariato. I dati sono espressi come coefficiente Beta standardizzato ( $\beta$ ), errore standard (SE).

Il modello di COX è stato implementato per stimare l'associazione tra i singoli Biomarcatori dello stato infiammatorio o lo score infiammatorio (suddiviso in terzili) e gli eventi vascolari maggiori (end-point combinato tra eventi cardiovascolari, eventi cerebrovascolari e periferici). Sono stati implementati due modelli, il primo non aggiustato ed il secondo aggiustato per i possibili confondenti Fumo, Ipertensione, Ipercolesterolemia, Diabete, Età, Sesso e Latitudine). I risultati sono presentati come Hazard Ratio (HR) e intervallo di confidenza al 95% (IC 95%).

Per valutare la capacità predittiva di evento cardiovascolare da parte dello score infiammatorio, confrontato con i singoli biomarcatori infiammatori o con i fattori di rischio cardiovascolare (Fumo, Ipertensione, Ipercolesterolemia, Diabete, Età, Sesso) più Latitudine è stata applicata una regressione logistica ed è stata calcolata l'area sotto la curva (AUC) delle curve ROC.

Il T- test è stato impiegato per valutare la differenza dei valori medi dello score nella popolazione esente da eventi e nei pazienti che hanno subito un evento vascolare maggiore.

Inoltre, è stata implementata la curva delle incidenze di Kaplan-Meier per valutare l'associazione tra i terzili dello score infiammatorio e gli eventi vascolari maggiori.

## **Score di stato infiammatorio**

È stato creato uno score che rifletta lo stato di infiammazione dei pazienti dello studio IMPROVE, utilizzando il pannello di 4 biomarcatori dello stato di infiammazione (Globuli Bianchi, Interleuchina 6, sgp130 e Proteina c-reattiva), combinati con pesi proporzionali alla loro associazione multivariata con l'aterosclerosi.

1° step: Regressione multivariata per studiare l'associazione tra i 4 biomarcatori di stato infiammatorio presi in considerazione e l'aterosclerosi subclinica, misurata con c-IMT basale. I coefficienti della regressione sono stati impiegati come pesi di ciascun biomarcatore da introdurre nello score.

2° step: creazione dello score di stato infiammatorio come segue:

$$\begin{aligned} \text{Score stato infiammatorio} = & -0.04 * \text{CRP} + \\ & 0.13 * \text{IL-6} + \\ & -0.03 * \text{sgp130} + \\ & 0.05 * \text{Globuli Bianchi}. \end{aligned}$$

## **Path Analysis e SEM**

Infine sono stati implementati modelli innovati quali la Path Analysis e i SEM per valutare gli effetti dei fattori di rischio 'classici', lo score infiammatorio, c-IMT, la sua progressione e gli eventi vascolare maggiori simultaneamente. Sono stati valutati 6 modelli differenti che includono le seguenti variabili:

- 1° Modello: Eventi vascolari maggiori, c-IMT, Fattori di rischio
- 2° Modello: Eventi vascolari maggiori, Progressione del c-IMT, Fattori di rischio
- 3° Modello: Eventi vascolari maggiori, c-IMT, Fattori di rischio e score infiammatorio

- 4° Modello: Eventi vascolari maggiori, Progressione del c-IMT, Fattori di rischio e score infiammatorio.
- 5° Modello: Eventi vascolari maggiori, c-IMT, Progressione c-IMT, Fattori di rischio e score infiammatorio.
- 6°Modello: Eventi vascolari maggiori, c-IMT, Progressione c-IMT, Fattori di rischio, Biomarcatori dello stato infiammatorio e variabile latente 'Infiammazione'.

In alternativa allo score di stato infiammatorio è stata dunque ipotizzata una variabile latente denominata 'Infiammazione'. Tale variabile (non direttamente osservabile) rappresenta lo stato infiammatorio sistemico e riflette l'effetto congiunto dei 4 biomarcatori di stato di infiammazione. A differenza dello score di stato infiammatorio, che è stato costruito partendo dalle conoscenze a priori, nella scelta dei biomarcatori, e dalla regressione di questi ultimi con il c-IMT basale, la variabile latente è prodotta dalle equazioni strutturali, le quali tengono conto dell'intero insieme di relazioni esistenti tra tutte le variabili.

## **Indici di adattamento**

In letteratura esistono diversi indici di adattamento per stabilire se il modello stimato con la SEM o con la Path Analysis sono dei buoni modelli.

Questi indici si suddividono in:

- Indici di adattamento assoluti, che valutano la bontà di adattamento del modello ai dati
- Indici di adattamento di parsimonia, che valutano se, escludendo le variabili o le relazioni superflue, la capacità esplicativa del modello rimane elevata.

Gli indici di adattamento assoluti, comprendono: Goodness-of-fit (GFI), Root Mean Square Residual (RMSR) e Test Chi-quadro.

Il GFI assume valori in  $[0,1]$  e può essere interpretato in modo analogo all'indice di bontà del modello di regressione  $R^2$ ; quando il GFI assume valori non compresi nell'intervallo indicato, ciò significa che i dati non supportano la struttura alla base del modello Path/SEM. L'indice GFI aumenta all'aumentare della numerosità campionaria.

Il RMSR fornisce la media quadratica delle correlazioni residue non spiegate dal modello PATH, per cui, con valori prossimi a zero ( $<0.05$ ) sta ad indicare un buon adattamento del modello.

Test del Chi-quadro: nel caso di perfetto adattamento dei dati, assumerà valori uguali a zero. Nel nostro studio non lo useremo perché è troppo dipendente dalla numerosità campionaria.

Gli indici di parsimonia, invece, comprendono: Adjusted GFI (AGFI), Indice di stima comparativo di Bentler (CFI). Questi indici, a differenza dei precedenti, prendono in considerazione la comprensività del modello.

AGFI, è un aggiustamento dell'indice GFI, per i gradi di libertà del modello. L'intervallo non è più in  $[0,1]$  e in caso di adattamento perfetto, assume il valore 1; valori maggiori di 0.9 indicano un buon adattamento del modello.

CFI, confronta il modello selezionato con il modello definito 'base', ovvero quello in cui tutte le covariate assumono valori uguali a zero. Valori di  $CFI > 0.9$  significano un buon grado di adattamento del modello.

Tutti i test sono stati condotti su due code ed eseguiti dal software statistico SAS v. 9.4 (SAS Institute Inc., Cary, NC, USA). La Path Analysis e i SEM sono stati implementati utilizzando la procedura Proc CALIS del software SAS.

# Risultati

## Caratteristiche della popolazione

Le caratteristiche di base dei partecipanti sono riportate nella tabella 1.

L'età media era di 64,2 anni e il 47,9% dei soggetti erano maschi. I pazienti risultavano avere una tendenza al sovrappeso, dato che mediamente hanno un BMI di 27.3 Kg/m<sup>2</sup>, cioè superiore al cut-off 25, oltre il quale si è valutati Obesi di I categoria. (De Lorenzo, Soldati et al. 2016)

Inoltre, circa il 20% dei soggetti può essere considerato sedentario dato che dichiara di svolgere bassa attività fisica (una camminata di almeno 10 minuti meno di una volta alla settimana). Quasi il 50% dei pazienti (48.2%) ha dichiarato di non aver mai fumato. Il valore mediano di 22.5 del Framingham risk score conferma che i pazienti reclutati erano soggetti ad alto rischio. I pazienti sono stati seguiti per un follow-up mediano di 21,5 mesi, periodo in cui si sono riscontrati 215 eventi vascolari (125 cardiovascolari, 73 cerebrovascolari e 17 periferici).

Tabella 1: Caratteristiche di base dei pazienti partecipanti

<b>Centri</b>		
Milano, n(%)		553 (14.93)
Kuopio, n(%)		1048 (28.3)
Stockholm, n(%)		532 (14.37)
Groeningen, n(%)		527 (14.23)
Perugia, n(%)		542 (14.64)
Paris, n(%)		501 (13.53)
<b>Varibili Antropometriche</b>		
Maschi, n(%)		1774(47.91)
Età ( anni)		64.17±5.44
BMI (KG/m2)		27.27±4.27
Pressione sistolica (mmHg)		141.97±18.49
Pressione diastolica (mmHg)		81.96±9.79
Anni di studio		10.43±3.9
Attività fisica		
	Bassa, n(%)	731 (19.83)
	Media, n(%)	1635 (44.36)
	Intensa , n(%)	1320 (35.81)
C-IMT (mm)		1.42±0.41
Follow-up (mesi)		36.7(36.3-37.9)
<b>Abitudine al fumo</b>		
Fumatore		549 (14.83)
Ex-fumatore		1371(37.02)
Mai fumato		1783 (48.15)
PackYears		18 (8-30)
<b>Fattori di rischio</b>		
Colesterolo Totale (mg/dl)		211.97±43.61
Trigliceridi (mg/dl)		114.22 (81.71-166.06)
HDL (mg/dl)		211.97±43.61
LDL (mg/dl)		134.85(108.74-162.50)
Glucosio		136.56±38.82
Creatinina		79.10(68.40-90.70)
Framingham risk score		22.55 (14.31 - 34.35)
<b>Biomarcatori dello stato infiammatorio</b>		
Proteina c-Reattiva (mg/l)		1.85(0.76-3.58)
Interleuchina 6(ng/ml)		30.25(22.02-46.25)
sgp130 (ng/ml)		589.8 ±181.03
Globuli Bianchi		6.14 ± 1.76

**Trattamento farmacologico**

Statine, n(%)	1483 (40.05)
Fibrati, n(%)	284 (7.67)
Antipertensivi, n(%)	2109 (56.95)
Beta Bloccanti, n(%)	878 (23.71)
Calcio Antagonista, n(%)	603 (16.28)
ACE inebitori, n(%)	722 (19.5)
Diuretici, n(%)	857 (23.14)
Insulina, n(%)	141(3.81)

**Familiarità**

Familiarità CHD, n(%)	2315(64.8)
Familiarità CVD, n(%)	1322(35.7)
Familiarità PVD, n(%)	443(11.9)

**Eventi**

Eventi combinati	215 (5.81)
Eventi cardiovascolari	125 (3.38)
Eventi Cerebrovascolari	73 (1.97)
Eventi Periferisci	17 (0.46)

---

*I dati sono rappresentanti come media  $\pm$  deviazione standard o mediana (1° quartile- 3° quartile) per le variabili continue, frequenza (percentuale) per le variabili categoriche. 'Packyears' è stata calcolata come pacchetti di sigarette fumate mediamente in un giorno \* il numero di anni in cui il soggetto ha fumato; è stata calcolata escludendo i non fumatori. L'attività fisica è stata definita come 'Bassa' per i soggetti che dichiarano di effettuare una camminata veloce di almeno 10 min < 1 volta a settimana, 'Media' per i soggetti che svolgono attività almeno 2/3 volte a settimana ed 'Intensa' per i soggetti che svolgono attività più di 2/3 volte a settimana.*

## Relazione c-IMT e Biomarcatori dello stato infiammatorio

### Correlazione reciproca dei Biomarcatori dello stato infiammatorio

I biomarcatori di stato infiammatorio presi in considerazione per le analisi sono: Proteina C-Reattiva (CRP), Interleuchina-6 (IL-6), sgp130 e Globuli Bianchi. L'IL-6 e la CRP correlano abbastanza bene tra di loro ( $r= 0.42$ ,  $p\text{-value} < 0.0001$ ). La correlazione osservata tra Globuli Bianchi e CRP e tra Globuli Bianchi e IL-6 è significativa, ma relativamente modesta ( $r=0.26$  e  $r=0.28$ , rispettivamente). Si osserva, infine, la scarsa correlazione tra sgp130 e CRP ( $r=0.06$ ) e l'assenza di correlazione significativa tra sgp130 e globuli bianchi e tra CRP e IL-6 ( $r=0.02$ ). [Tabella 2]

Tabella 2: Correlazioni tra Biomarcatori dello stato di infiammazione

	Proteina c-Reattiva (mg/l)	Interleuchina- 6	sgp130	Globuli Bianchi
<i>Proteina c-Reattiva (mg/l)</i>	-	0.42 (<.0001)	0.06 (0.0009)	0.26 (<.0001)
<i>Interleuchina-6</i>		-	0.02 (0.2839)	0.28 (<.0001)
<i>sgp130</i>			-	0.02 (0.3966)

*I dati sono espressi come coefficiente di correlazione, r, (p-value). La Proteina C-reattiva e l'Interleuchina-6 sono espresse in logaritmo.*

## Analisi con modelli di regressione semplice e multipla della relazione tra i Biomarcatori dello stato infiammatorio e c-IMT.

Nell'analisi univariata (Modello 1), tutti i biomarcatori di stato infiammatorio presi in considerazione, ad eccezione della CRP, sono associati al c-IMT basale. [Tabella 3].

Implementando una regressione multipla (Modello 2) che comprende tutti e quattro i biomarcatori nello stesso modello, sgp130 perde la sua significatività, mentre la CRP diviene significativa e, curiosamente, il coefficiente cambia di segno (esempio del cosiddetto 'paradosso di Simpson').

Tabella 3: Associazione tra i Biomarcatori dello stato infiammatorio e c-IMT

	$\beta$	SE	p-value
<b>Modello 1</b>			
<i>Proteina c-Reattiva (mg/l)</i>	0.02	0.004	0.2873
<i>Interleuchina-6</i>	0.12	0.003	<.0001
<i>sgp130</i>	-0.04	0.0001	0.023
<i>Globuli Bianchi</i>	0.07	0.002	<.0001
<b>Modello 2</b>			
<i>Proteina c-Reattiva (mg/l)</i>	-0.04	0.004	0.0168
<i>Interleuchina-6</i>	0.13	0.003	<.0001
<i>sgp130</i>	-0.03	0.0001	0.0746
<i>Globuli Bianchi</i>	0.05	0.001	0.006

*I Beta ( $\beta$ ) sono standardizzati. Il modello 1 non è aggiustato; nel Modello2, i biomarcatori sono aggiustati reciprocamente. La Proteina C-reattiva e l'Interleuchina-6 sono espresse in logaritmo.*

## Associazione eventi e biomarcatori dello stato infiammatorio: modelli di regressione.

La tabella 4 mostra le associazioni tra i diversi biomarcatori di stato d'infiammazione e l'evento vascolare combinato.

All'analisi univariata (Modello 1), ad eccezione del sgp130, tutti i Biomarcatori di stato d'infiammazione sono associati all'evento combinato (Tabella 4); aggiustando per i possibili confondenti [diabete, fumo, ipertensione, ipercolesterolemia, età, sesso e latitudine] (Modello 2), solo l'Interleuchina-6 resta un predittore indipendente degli eventi vascolari [OR (IC 95%): 1.70 (1.51;1.91)]

Tabella 4: Associazione tra i Biomarcatori di stato infiammatorio e l'evento vascolare combinato

	HR	IC 95%	p-value
<b>Modello 1</b>			
<i>Proteina c-Reattiva (mg/l)</i>	1.46	(1.12;1.90)	0.0054
<i>Interleuchina-6</i>	1.82	(1.63;2.02)	<.0001
<i>sgp130</i>	0.86	(0.56;1.33)	0.5033
<i>Globuli Bianchi</i>	1.74	(1.04;2.91)	0.0340
<b>Modello 2</b>			
<i>Proteina c-Reattiva (mg/l)</i>	1.26	(0.95;1.67)	0.1043
<i>Interleuchina-6</i>	1.70	(1.51;1.91)	<.0001
<i>sgp130</i>	0.96	(0.61;1.5)	0.8457
<i>Globuli Bianchi</i>	0.86	(0.47;1.58)	0.6388

HR: Hazard Ratio; IC95%: Intervallo di confidenza al 95%. Modello 1 non aggiustato. Modello 2 aggiustato per diabete, fumo, ipertensione, ipercolesterolemia, età, sesso e latitudine. La Proteina C-reattiva e l'Interleuchina-6 sono espresse in logaritmo.

## Fattori di rischio, c-IMT, eventi e score dello stato infiammatorio.

### Relazione tra score di stato di infiammazione e fattori di rischio, c-IMT

La tabella 5 mostra le correlazioni semplici tra i fattori di rischio (Diabete, Fumo, Ipercolesterolemia, età, sesso e latitudine) e lo score di stato infiammatorio, stimato come descritto prima (vedi metodi). Da sottolineare la correlazione positiva con c-IMT ( $r=0.14$ ,  $p\text{-value}<0.0001$ ).

*Tabella 5: Correlazioni tra i fattori di rischio, C-IMT e Score di stato infiammatorio*

	SCORE Infiammatorio
<i>c-IMT</i>	0.12 (<.0001)
<i>PackYears</i>	0.15 (<.0001)
<i>Ipercolesterolemia</i>	0.06 (0.003)
<i>Diabete</i>	0.09 (<.0001)
<i>Ipercolesterolemia</i>	-0.07 (<.0001)
<i>Età</i>	0.07 (<.0001)
<i>Sesso</i>	0.09 (<.0001)
<i>Latitudine</i>	0.13 (<.0001)

*I dati sono espressi come coefficiente di correlazione, r, (p-value).*

## Capacità predittiva degli eventi da parte dello score di stato di infiammazione: confronto con i singoli biomarcatori

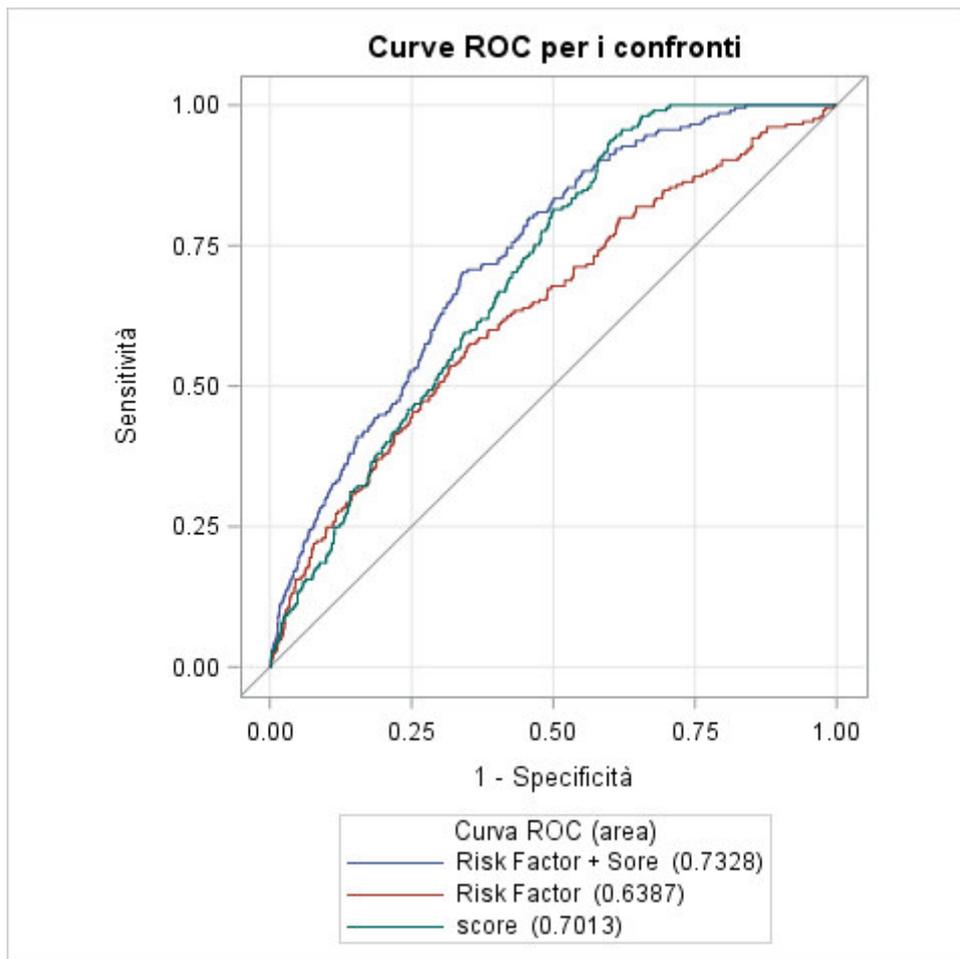
L'area sotto la curva dello score è più alta (AUC=0.70) rispetto le singole AUC dei singoli biomarcatori dello stato infiammatorio. [Tabella 6]

*Tabella 6: Area sotto la curva dello Score e dei singoli biomarcatori dello stato infiammatorio*

	AUC	IC 95%	
<i>Score</i>	0.7013	0.6719	0.7307
<i>PCR</i>	0.561	0.5203	0.6016
<i>sgp130</i>	0.5139	0.4741	0.5538
<i>Globuli Bianchi</i>	0.5531	0.5125	0.5937
<i>IL-6</i>	0.6922	0.6627	0.7217

AUC: Area sotto la curva. IC 95% : Intervallo di confidenza al 95%.

Lo score infiammatorio, inoltre, migliora in modo significativo il valore prognostico dei classici fattori di rischio (  $AUC_{\text{score+fattoridirischio}} = 0.73$  ,  $AUC_{\text{fattoridirischio}} = 0.64$  ; p-value 0.004). [FIGURA 7]



*Figura 7: Curva Roc dei singoli fattori di rischio e dello score di stato di infiammazione*

## Relazione tra score di stato di infiammazione ed eventi

### vascolari maggiori

Lo score di stato infiammatorio era mediamente più alto nei pazienti che hanno avuto l'evento rispetto al gruppo esente da evento ( $0.27 \pm 0.07$  vs.  $0.21 \pm 0.09$ , rispettivamente;  $p\text{-value} < .0001$ ).

Suddividendo lo score infiammatorio in terzili, dalla Figura 8, si può osservare un chiaro trend tra le curve di incidenza nei tre terzili.

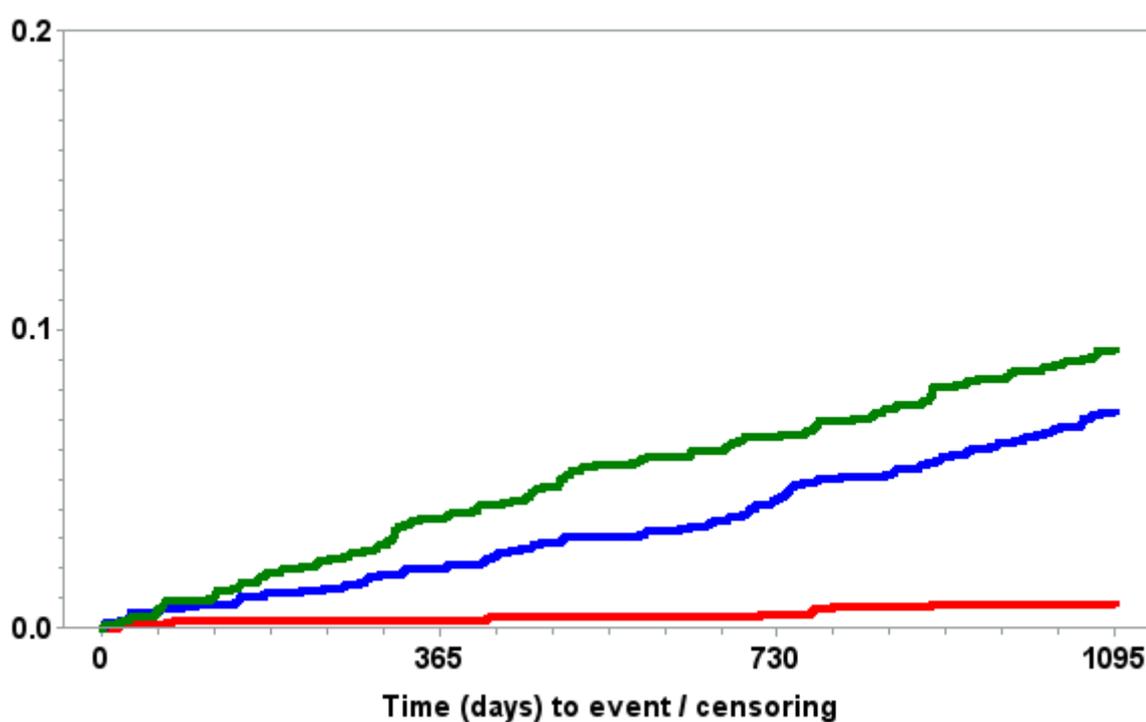


Figura 8: Curve di incidenza del evento vascolare combinato nei tre terzili di Score di stato di infiammazione. Linea rossa rappresenta il primo terzile, la linea blu il secondo terzile e la linea verde il terzo terzile.

Abbiamo osservato un trend significativo (HR (95%): 2.34 (1.93-2.84) per un incremento di terzile di score, p-value <.0001). Più precisamente, HR (CI 95%) 2° terzile vs 1° terzile: 9.1 (4.7-17.5); HR (CI 95%) 3° terzile vs 1° terzile : 12.2(6.4-23.3); HR (CI 95%) 2° terzile vs 3° terzile : 1.3 (1.1-1.8). [Figura 9]

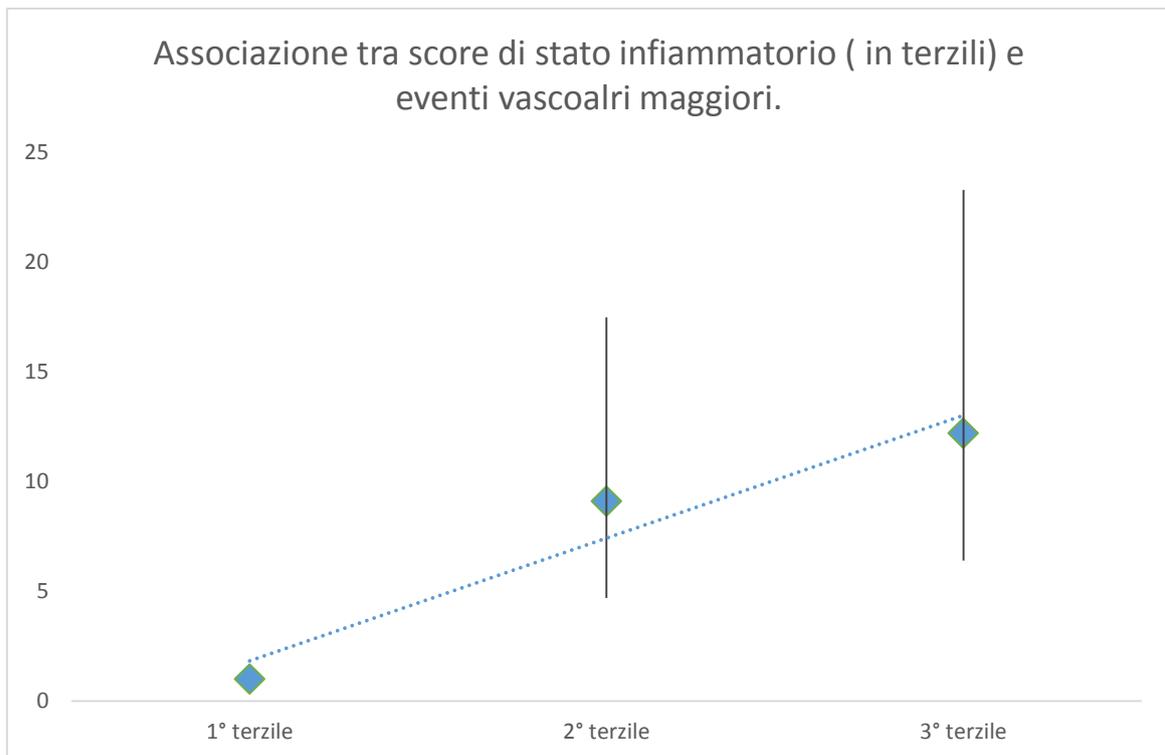


Figura 9: Associazione tra score di stato di infiammazione in terzili ed eventi vascolari maggiori. Hazard ratio (95% IC). Riferimento primo terzile.

Anche dopo aggiustamento per i fattori di rischio (Diabete, Fumo, Iperensione, Ipercolesterolemia, età, sesso e latitudine) lo score infiammatorio resta significativamente associato agli eventi [HR(CI 95%): 2.1(1.8-2.6); p-value <.0001]. Aggiungendo al modello anche c-IMT, il risultato resta invariato [HR(CI 95%): 2.1(1.7-2.6); p-value <.0001].

# Associazione tra eventi, c-IMT e fattori di rischio: Path Analysis e SEM

## Associazione eventi, c-IMT e fattori di rischio

Come anticipato nei capitoli precedenti, le relazioni osservate fino ad ora non tengono conto delle relazioni reciproche esistenti tra i singoli fattori predittivi. Abbiamo, quindi, implementato la Path Analysis e i SEM. Inizialmente abbiamo valutato le relazioni tra i diversi fattori di rischio: diabete, fumo, ipertensione, ipercolesterolemia, età, sesso e latitudine, variabili socio-economiche ('LAVORO'), l'attività fisica, c-IMT ed eventi vascolari maggiori.

La Figura 10 presenta le relazioni significative del modello sopradescritto:

- ➔ c-IMT, attività fisica e fumo sono associati agli eventi.
- ➔ Fattori di rischio come Sesso, età, ipertensione, ipercolesterolemia, latitudine, Professioni non sono direttamente associate agli eventi, ma al c-IMT. Sembra dunque che la loro relazione finale con gli eventi sia mediata dalla loro relazione con c-IMT.
- ➔ Variabili come Lavoro, Sesso, Latitudine, Fumo, Attività fisica ed età sono correlate ai fattori di rischio quali diabete, ipertensione e ipercolesterolemia.

I coefficienti dei singoli effetti diretti e indiretti sono descritti nella Tabella 8.

Gli indici di adattamento (Tabella 7) dimostrano, essendo SRMR < 0.05 e AGFI, CFI > 0.9, che il modello descritto presenta un buon grado di adattamento.

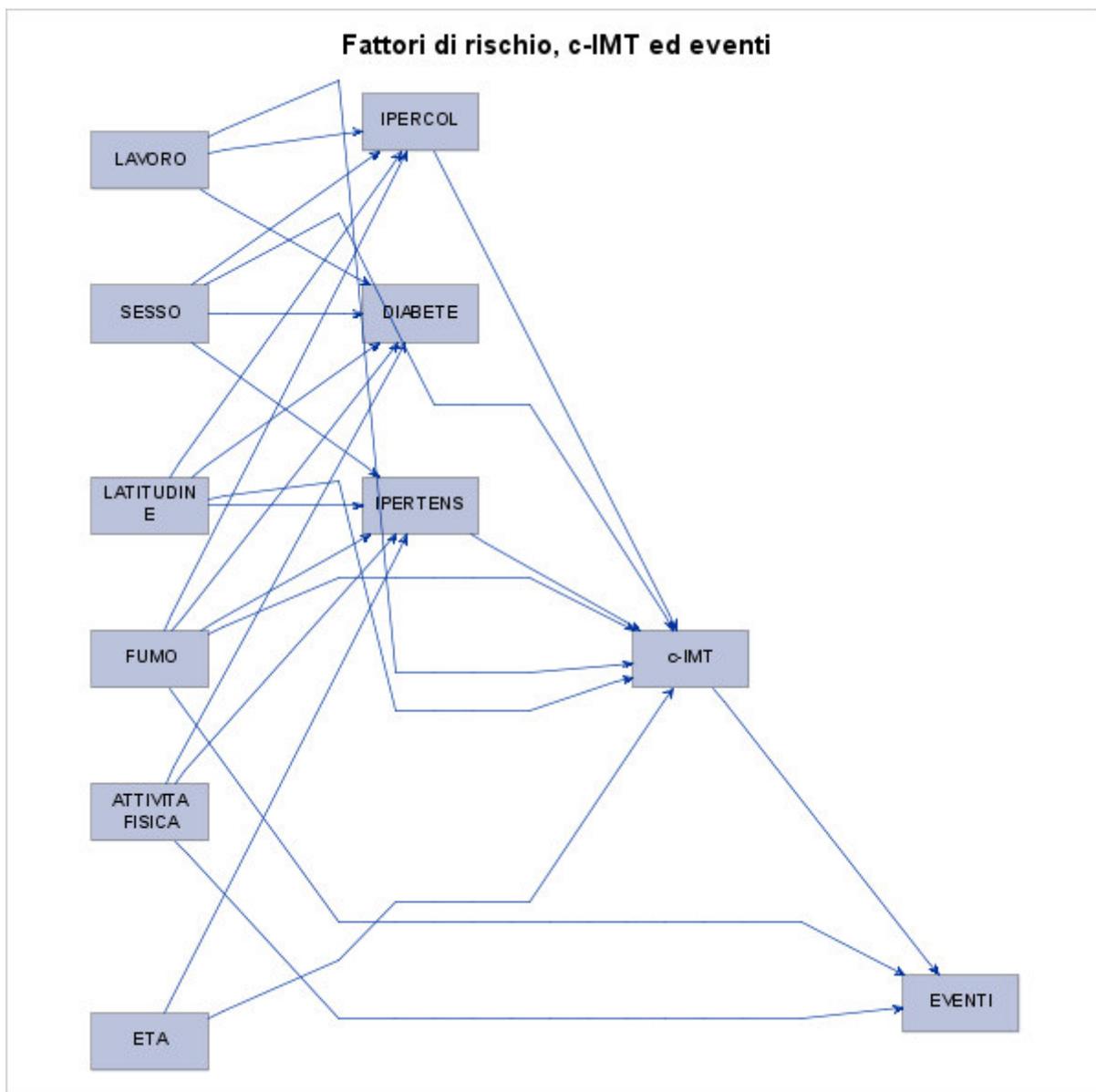


Figura 10: Path Diagram con effetti diretti di Fattori di rischio, c-IMT ed eventi vascolari maggiori

Tabella 7: Indici di Bontà di Adattamento del modello

INDICI DI BONTA' DI ADATTAMENTO	
RMR standardizzato (SRMR)	0.0275
Adjusted GFI (AGFI)	0.9994
Indice di stima comparativo di Bentler (CFI)	0.9221

Tabella 8: Coefficienti degli effetti diretti ed indiretti

Effetti diretti										
	Diabete	Ipercolesterolemia	Ipertensione	c-IMT	Età	Sesso	Professioni	Latitudine	Fumo	Attività fisica
<i>Diabete</i>	-	-	-	-	-0.76	4.01*	3*	7.52*	5.01*	-2.8*
<i>Eventi</i>	0.04	-0.91	0.48	4.29*	1.32	1.54	0.24	0.65	2.97*	-2.07*
<i>Ipercolesterolemia</i>	-	-	-	-	-1.75	5.09*	-5*	-13.19*	4.78*	1.71
<i>Ipertensione</i>	-	-	-	-	4.79*	2.08*	1.5	17.91*	3.43*	-2.36*
<i>c-IMT</i>	1.2	2.15*	6.53*	-	13.56*	9.9*	5.55*	11.01*	8.16*	0.41
Effetti indiretti										
<i>Diabete</i>	-	-	-	-	-	-	-	-	-	-
<i>Eventi</i>	1.15	1.92	3.59*	-	4.19*	3.45*	2.96*	2.85*	2.91*	-0.34
<i>Ipercolesterolemia</i>	-	-	-	-	-	-	-	-	-	-
<i>Ipertensione</i>	-	-	-	-	-	-	-	-	-	-
<i>c-IMT</i>	-	-	-	-	3.03*	2.05*	0.21	3.60*	2.51*	-1.85

\*gli effetti statisticamente significativi (p-value <0.05)

## **Associazione tra eventi, progressione del c-IMT e fattori di rischio**

Abbiamo implementato un secondo modello, sostituendo la c-IMT con la sua progressione, valutata mediante la 'Fastest-IMTmax-progression'. Si evidenziano, così, le relazioni con la velocità di progressione di malattia, anziché con la condizione basale dei pazienti.

Introducendo nel modello la progressione della c-IMT anziché la c-IMT, notiamo che in questo caso è la progressione della malattia ad essere direttamente associata agli eventi. [Figura 11]

I fattori che hanno un effetto diretto (vedi i coefficienti nella tabella 10) verso la progressione c-IMT sono:

- Età
- Fumo
- Latitudine
- Sesso
- Diabete

Ad eccezione dell'età, tutte le variabili, influenzano anche ipercolesterolemia e diabete. L'ipertensione, invece, è influenzata dall'attività fisica, età, fumo e latitudine.

Due indici su tre (SRMR e AGFI) soddisfano i requisiti per indicare un buon adattamento del modello. Il CFI è al limite, avendo un valore di 0.8947 (per essere soddisfacente dovrebbe superare 0.9). [Tabella 9]

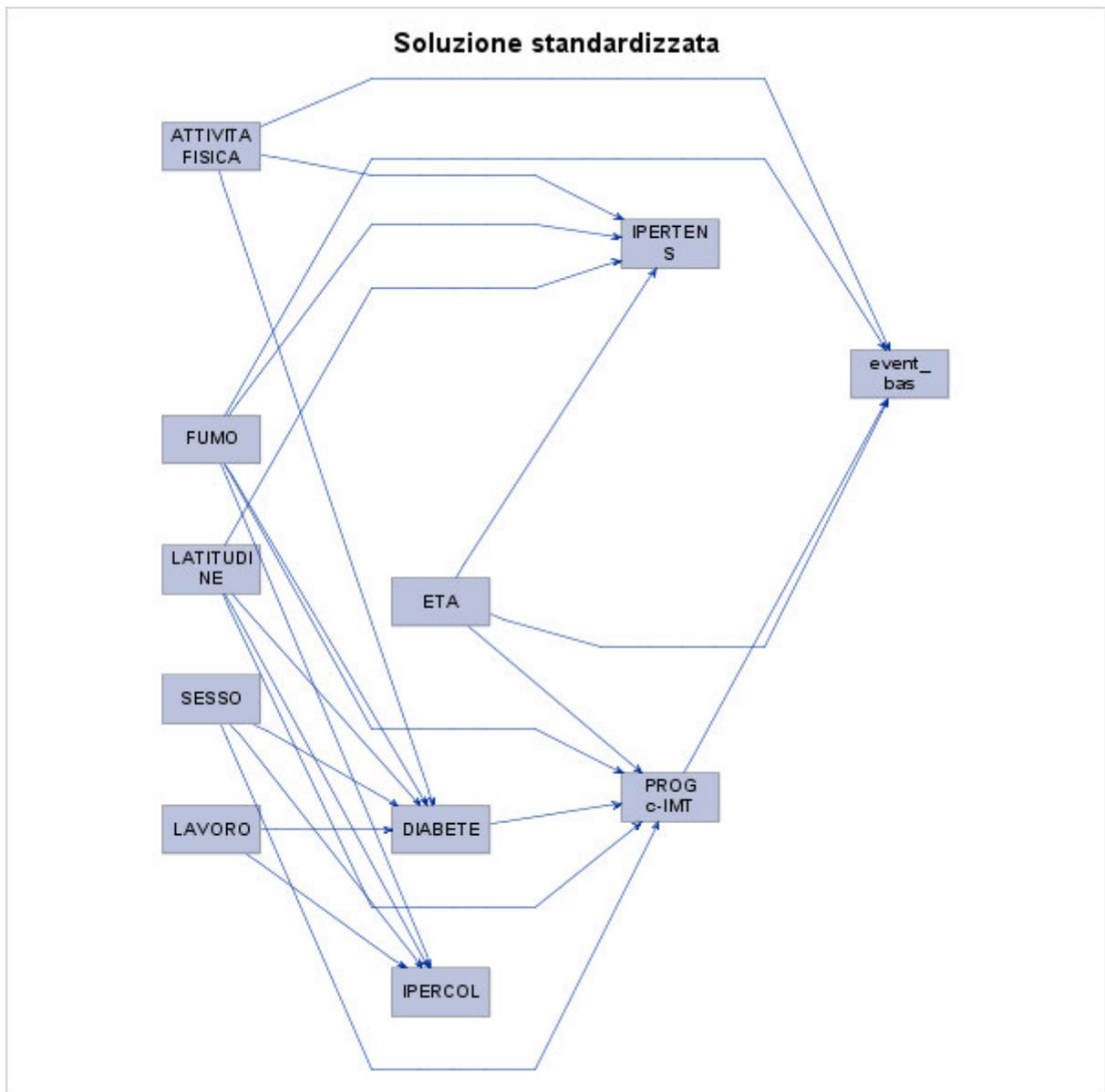


Figura 11: Path Diagram con effetti diretti Fattori di rischio, progressione c-IMT ed eventi vascolari maggiori

Tabella 9: Indici di Bontà di adattamento del modello

INDICI DI BONTA' DI ADATTAMENTO	
RMR standardizzato (SRMR)	0.0286
Adjusted GFI (AGFI)	0.9980
Indice di stima comparativo di Bentler (CFI)	0.8947

Tabella 10: Coefficienti diretti ed indiretti

	Effetti diretti									
	Diabete	Ipercolesterolemia	Ipertensione	Progressione c-IMT	Età	Sesso	Professioni	Latitudine	Fumo	Attività fisica
Diabete	-	-	-	-	-0.45	3.50*	2.633*	7.02*	5.09*	-2.74*
Eventi	-0.06	-0.49	0.66	2.42*	2.37	1.88	0.47	1.78	3.13*	-2.22*
Ipercolesterolemia	-	-	-	-	-1.35	-4.54*	-4.48*	-12.86*	-4.86*	1.38
Ipertensione	-	-	-	-	4.10*	-1.78	1.17	17.25*	-3.44*	-1.98*
Progressione c-IMT	2.85*	1.81	1.41	-	3.8*	2.22*	0.13	4.46*	3.90*	-1.51
	Effetti indiretti									
Diabete	-	-	-	-	-	-	-	-	-	-
Eventi	1.86	1.41	1.18	-	1.97*	0.84	0.46	1.16	0.96	-1.08
Ipercolesterolemia	-	-	-	-	-	-	-	-	-	-
Ipertensione	-	-	-	-	-	-	-	-	-	-
Progressione c-IMT	-	-	-	-	0.34	-0.26	0.05	0.78	-0.02	-1.52

\*Gli effetti statisticamente significativi (p-value <0.05)

## **Associazione tra eventi, c-IMT, fattori di rischio e score dello stato infiammatorio.**

Introducendo lo score dello stato infiammatorio che abbiamo stimato nel modo descritto precedentemente, notiamo dalla figura 12, che tale score assume una posizione abbastanza centrale nel grafico.

- ➔ Lo score è associato significativamente sia agli eventi, sia al c-IMT;
- ➔ Lo score è influenzato direttamente dalle seguenti variabili: sesso, latitudine, fumo, professione ed età.

con l'introduzione dello score nel modello, il fumo e l'attività fisica perdono la loro associazione con gli eventi, osservata nella figura 12.

Da notare che le relazioni tra i fattori intermedi (diabete, ipertensione, ipercolesterolemia e score di stato infiammatorio) non sono state valutate.

I valori dei coefficienti degli effetti diretti e indiretti sono visibili nella Tabella 12.

Tutte e tre gli indici soddisfano i criteri di un buon adattamento del modello. [TABELLA 11]

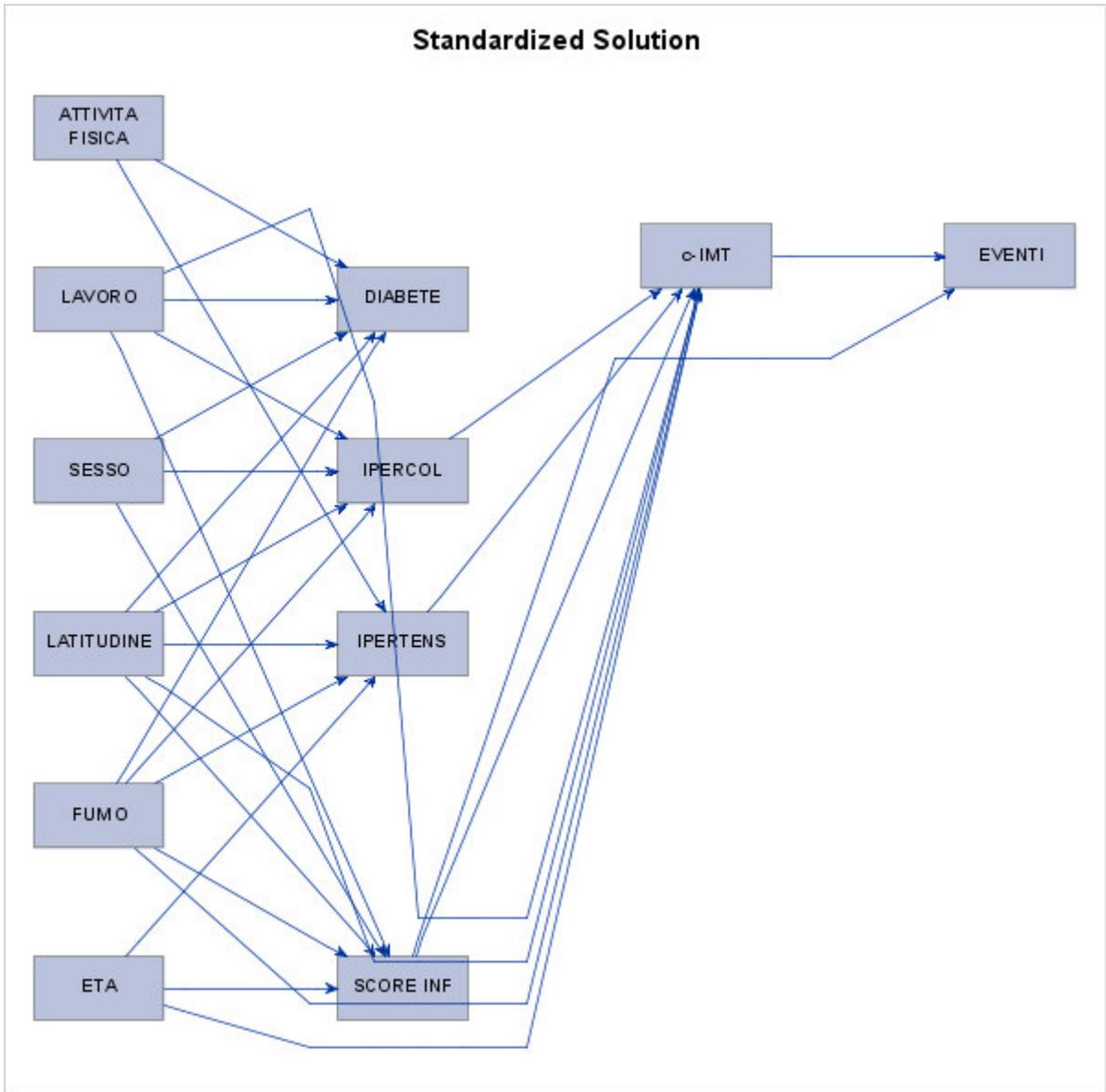


Figura 12: Path Diagram con effetti diretti fattori di rischio, c-IMT, score di stato di infiammazione ed eventi vascolari maggiori

Tabella 11: Indici di Bontà di adattamento del modello

---

INDICI DI BONTA' DI ADATTAMENTO	
RMR standardizzato (SRMR)	0.0262
Adjusted GFI (AGFI)	0.99899
Indice di stima comparativo di Bentler (CFI)	0.9224

---

Tabella 12: Coefficienti degli effetti diretti ed indiretti

Effetti diretti												
	Diabete	Ipercolesterolemia	Iperensione	C- IMT	Score infiammatorio	Età	Sesso	Professioni	Latitudine	Fumo	Attività fisica	
Diabete	-	-	-	-	-	-0.68	3.67*	3.266*	7.61*	5.23*	-2.30*	
Eventi	-0.81	-0.71	0.35	4.19*	9.30*	1.07	0.98	0.2	0.07	1.93	-1.64	
Ipercolesterolemia	-	-	-	-	-	-1.34	5.51*	-4.89*	-12.72*	3.91*	1.43	
Iperensione	-	-	-	-	-	3.79*	-1.79	1.58	17.37*	3.82*	-2.03*	
c-IMT	1.94	1.66	6.16*	-	1.94*	13.42*	9.15*	5.66*	9.84*	8.01*	0.87	
Score Infiammatorio	-	-	-	-	-	2.45*	2.01*	2.12*	6.85*	6.67*	-1.03	
Effetti indiretti												
	Diabete	Ipercolesterolemia	Iperensione	C- IMT	Score infiammatorio	Età	Sesso	Professioni	Latitudine	Fumo	Attività fisica	
Diabete	-	-	-	-	-	-	-	-	-	-	-	
Eventi	1.76	1.54	3.48*	-	1.76	4.73*	3.56*	3.28*	3.90*	5.07*	-0.50	
Ipercolesterolemia	-	-	-	-	-	-	-	-	-	-	-	
Iperensione	-	-	-	-	-	-	-	-	-	-	-	
C-IMT	-	-	-	-	-	2.65*	-1.27	0.85	3.83*	-1.50	-1.97*	
Score Infiammatorio	-	-	-	-	-	-	-	-	-	-	-	

\*\_gli effetti statisticamente significativi (p-value <0.05)

## **Associazione tra eventi, progressione c-IMT, fattori di rischio e score di stato infiammatorio.**

Sostituendo, nell'ultimo modello sopra descritto il c-IMT con la sua progressione, notiamo che lo score infiammatorio mantiene un ruolo centrale. Inoltre possiamo osservare che:

- ➔ Lo score, età, progressione c-IMT e fumo sono direttamente associati agli eventi.
- ➔ Come nel modello precedente, sesso, latitudine, fumo ed età influenzano lo score, mentre la relazione con la professione perde di significatività.

Tutte le relazioni sono sintetizzate nella FIGURA 13 ed i coefficienti relativi alle relazioni dirette e indirette sono rappresentati nella TABELLA 14.

Gli indici di adattamento soddisfano quasi tutti i criteri, ad eccezione di CFI che è pari a 0.8993.

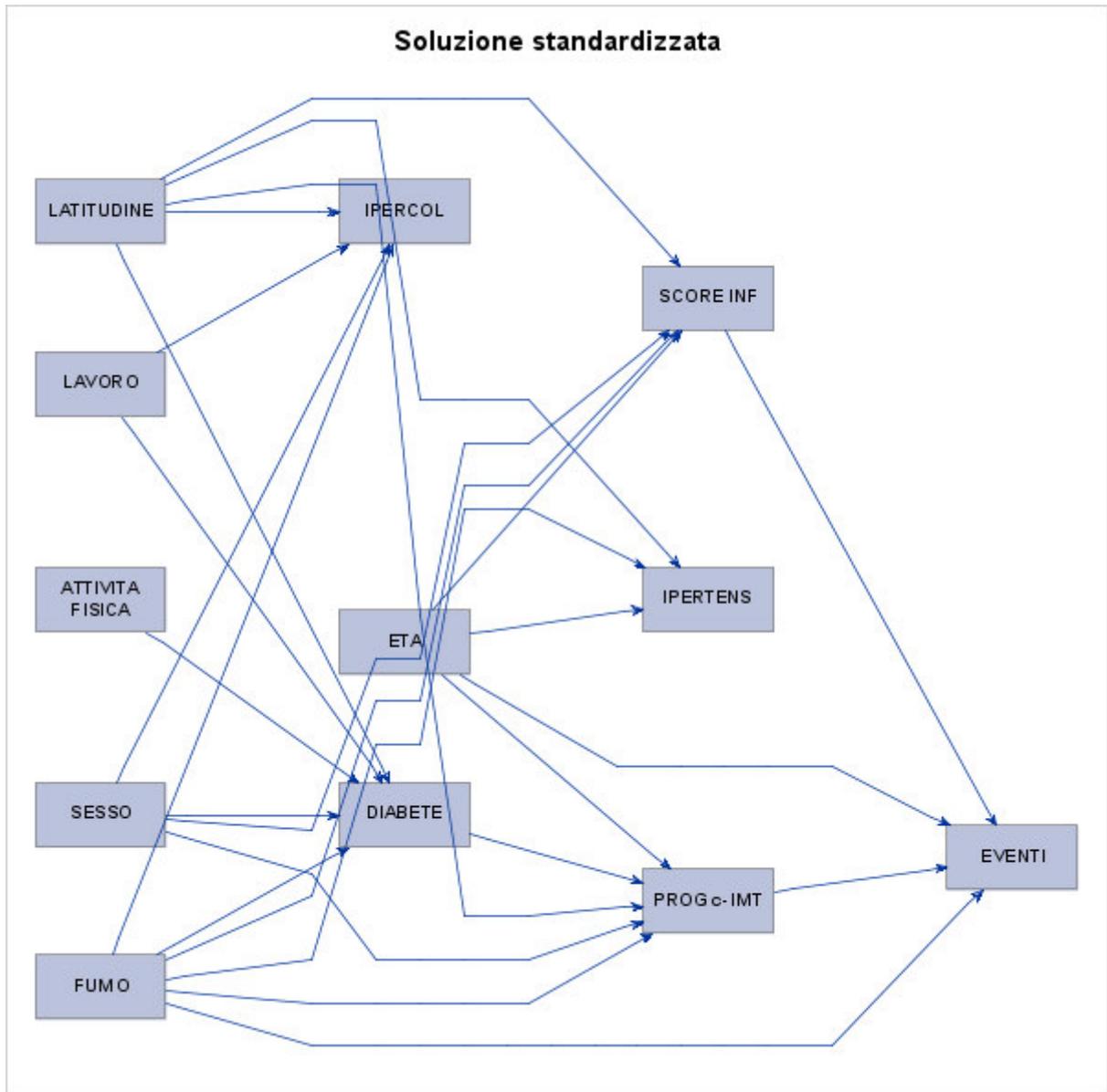


Figura 13: Path Diagram con effetti diretti Fattori di rischio, progressione della c-IMT, score dello stato infiammatorio ed eventi vascolari maggiori

*Tabella 13: Indici di Bontà di adattamento del modello*

---

INDICI DI BONTA' DI ADATTAMENTO	
RMR standardizzato (SRMR)	0.0267
Adjusted GFI (AGFI)	0.9983
Indice di stima comparativo di Bentler (CFI)	0.8993

---

Tabella 14: Coefficienti degli effetti diretti ed indiretti

Effetti diretti											
	Diabete	Ipercolesterolemia	Iperensione	Progressione c-IMT	Score infiammatorio	Età	Sesso	Professioni	Latitudine	Fumo	Attività fisica
Diabete	-	-	-	-	-	-0.25	3.28*	2.86*	7.13*	5.17*	-2.45*
Eventi	-0.65	-0.29	0.68	1.92*	9.43*	1.99*	1.4	0.41	1.22	2.18*	-1.82
Ipercolesterolemia	-	-	-	-	-	-0.94	4.83*	-4.38*	-12.36*	3.80*	1.22
Iperensione	-	-	-	-	-	3.19*	-1.67	1.14	16.60*	3.80*	-1.70
Progressione c-IMT	2.81*	1.45	1.41	-	0.98	4.23*	2.22*	0.06	3.90*	3.52*	-1.24
Score Infiammatorio	-	-	-	-	-	2.52*	2.34*	1.46	6.69*	6.82*	-1.60
Effetti indiretti											
	Diabete	Ipercolesterolemia	Iperensione	Progressione c-IMT	Score infiammatorio	Età	Sesso	Professioni	Latitudine	Fumo	Attività fisica
Diabete	-	-	-	-	-	-	-	-	-	-	-
Eventi	1.6	1.14	1.1	-	0.86	3.02*	1.8	1.09	2.55*	4.01*	-1.61
Ipercolesterolemia	-	-	-	-	-	-	-	-	-	-	-
Iperensione	-	-	-	-	-	-	-	-	-	-	-
Progressione c-IMT	-	-	-	-	-	0.88	0.14	0.57	1.21	0.88	-1.75
Score Infiammatorio	-	-	-	-	-	-	-	-	-	-	-

\*gli effetti statisticamente significativi (p-value <0.05)

## **Associazione tra eventi, c-IMT, progressione c-IMT, fattori di rischio e score di stato infiammatorio.**

Il passo successivo, è stato quello di implementare un modello che comprendesse sia c-IMT che progressione del c-IMT.

Focalizzandosi sullo score infiammatorio ('SCORE INF') notiamo che esso influenza direttamente l'evento e c-IMT, ed è influenzato da sesso, latitudine, fumo, età. Da notare che inserendo c-IMT, progressione del c-IMT e lo score nello stesso modello, la progressione c-IMT perde il suo effetto diretto nei confronti degli eventi. [FIGURA 14]

Come già anticipato, lo score di stato infiammatorio risulta avere un effetto diretto su c-IMT ed eventi vascolari maggiori, ma non sulla progressione del c-IMT.

La Tabella 16 sintetizza i coefficienti degli effetti diretti.

Anche in questo caso, tutti gli indici di adattamento supportano il modello. (Tabella 15)

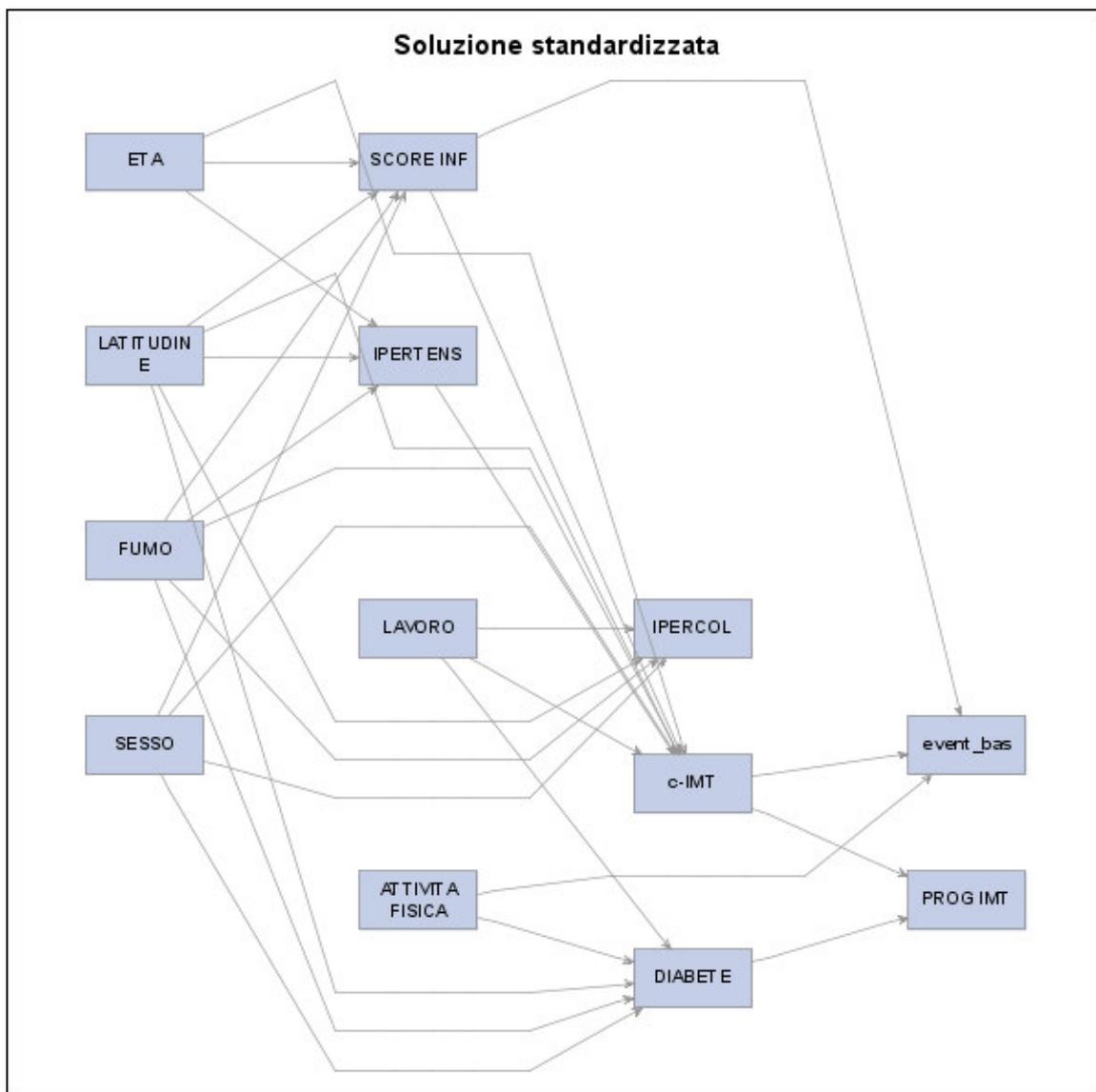


Figura 14: Path Diagram con effetti diretti. Fattori di rischio, C-IMT, Progressione del C-IMT, score dello stato infiammatorio ed eventi vascolari

Tabella 15: Indici di Bontà di adattamento del modello

INDICI DI BONTA' DI ADATTAMENTO	
RMR standardizzato (SRMR)	0.0249
Adjusted GFI (AGFI)	0.9982
Indice di stima comparativo di Bentler (CFI)	0.9233

Tabella 16: Coefficienti degli effetti diretti ed indiretti

Effetti diretti												
	Diabete	Ipercolesterolemia	Iperensione	Progressione c-IMT	Score infiammatorio	Età	Sesso	Professioni	Latitudine	Fumo	Attività fisica	
Diabete	-	-	-	-	-	-0,22	3,30*	2,82*	7,16*	5,18*	-2,48*	
Eventi	-0,70	-0,39	-0,17	1,04*	9,22*	1,06	0,68	-0,03	0,5	1,6	-1,96*	
Ipercolesterolemia	-	-	-	-	-	-0,96	4,80*	-4,39*	-12,37*	-	1,29	
Iperensione	-	-	-	-	-	3,18*	-1,68	1,15	16,56*	-	-1,70	
Progressione c-IMT	2,60*	1,15	0,13	-	0,48	1,42	0,37	-1,09	1,91	1,9	-1,54	
c-IMT	1,04	1,61	6,2	-	2,55*	13,32*	8,90*	5,42*	9,37*	7,75*	1,12	
Score Infiammatorio	-	-	-	-	-	2,49*	2,30*	1,5	6,64*	6,80*	-1,5	
Effetti indiretti												
	Diabete	Ipercolesterolemia	Iperensione	Progressione c-IMT	Score infiammatorio	Età	Sesso	Professioni	Latitudine	Fumo	Attività fisica	
Diabete	-	-	-	-	-	-	-	-	-	-	-	
Eventi	1,37	1,67	3,57*	-	2,24*	4,95*	3,83*	2,84*	3,86*	5,27*	-0,83	
Ipercolesterolemia	-	-	-	-	-	-	-	-	-	-	-	
Iperensione	-	-	-	-	-	-	-	-	-	-	-	
Progressione c-IMT	1,03	1,61	5,46*	-	2,49*	8,34*	6,20*	4,59*	5,14*	5,75*	0,29	
c-IMT	-	-	-	-	-	2,90*	-1,00	0,65	3,95*	-0,63	-1,87	
Score Infiammatorio	-	-	-	-	-	-	-	-	-	-	-	

\*gli effetti statisticamente significativi (p-value <0.05)

## **Associazione tra eventi, c-IMT, progressione c-IMT, fattori di rischio e variabile latente 'Infiammazione'.**

Infine, abbiamo implementato un modello SEM introducendo una nuova variabile, la variabile latente 'Infiammazione'. Questa variabile è stata costruita ipotizzando la relazione con i 4 biomarcatori di infiammazione ed è stata inserita nel modello comprendente i fattori di rischio, gli eventi vascolari maggiori, c-IMT, progressione di c-IMT ed i singoli biomarcatori di stato d'infiammazione.

Possiamo osservare (Figura 15) un effetto diretto di tale variabile latente verso gli eventi e il c-IMT. Attività fisica, Professioni, Fumo e latitudine, invece, influenzano la variabile latente.

Le frecce che collegano la variabile latente con interleuchina 6, Globuli Bianchi, Proteina C-Reattiva e sgp130 riflettono semplicemente il modo in cui è stata costruita la variabile latente, e non sono indici di un rapporto causale.

Anche in questo caso il grafico non mostra una relazione diretta tra progressione del C-IMT ed eventi.

I valori dei coefficienti sono sintetizzati nella tabella 18.

Anche in questo modello non tutti gli indici di adattamento portano a concludere che il modello di adatti bene ai dati, infatti, CFI è marcatamente minore di 0.9. (Tabella 17)

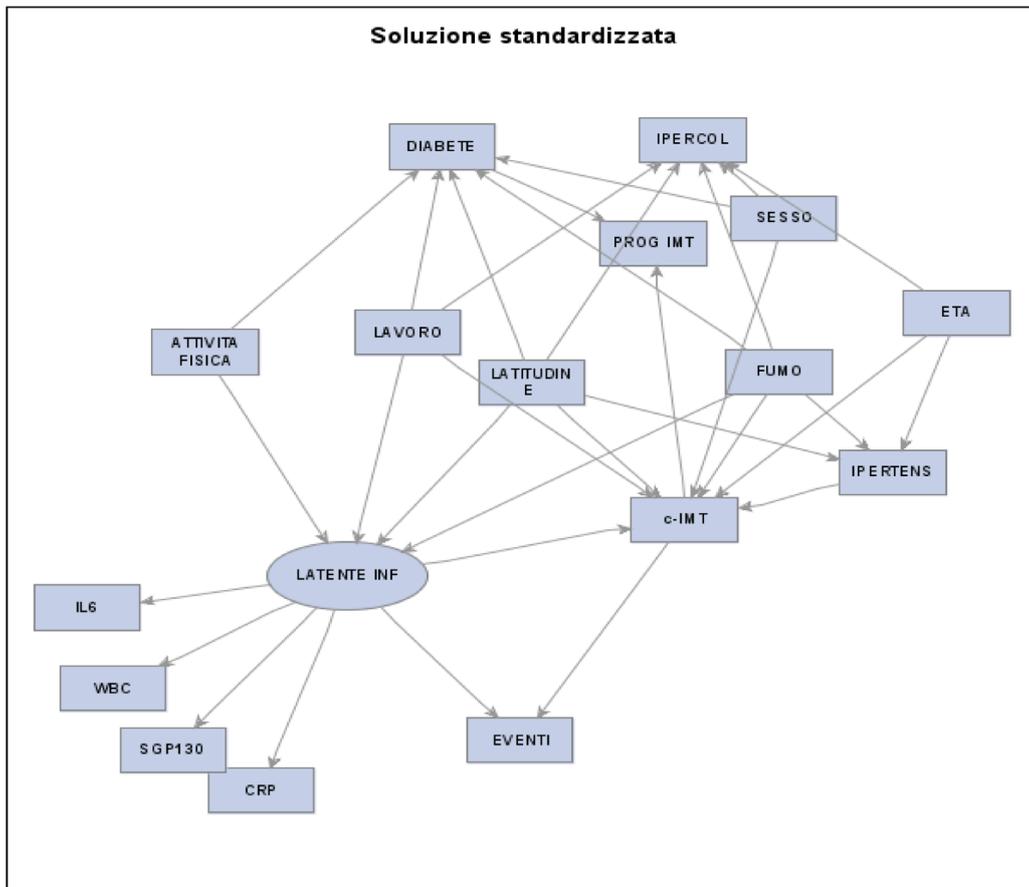


Figura 15: Path Diagram degli effetti diretti. Fattori di rischio, C-IMT, Progressione C-IMT, infiammazione come variabile latente ed eventi vascolari maggiori.

Tabella 17: Indici di Bontà di adattamento del modello

INDICI DI BONTA' DI ADATTAMENTO	
RMR standardizzato (SRMR)	0.0498
Adjusted GFI (AGFI)	0.9989
Indice di stima comparativo di Bentler(CFI)	0.8201

Tabella 18: Coefficienti degli effetti diretti ed indiretti

Effetti diretti											
	Diabete	Ipercolesterolemia	Iperensione	Progressione c-IMT	Età	Sesso	Professioni	Lattitudine	Fumo	Attività fisica	
Diabete	-	-	-	-	-0,24	3,27*	2,63*	7,42*	5,23*	-2,84*	
Eventi vascolari maggiori	-0,79	-0,36	0,06	0,03	-0,09	1,04	-0,4	-1,18	0,95	-1,52	
Ipercolesterolemia	-	-	-	-	-2,04*	-6,19*	-3,82*	-14,1*	-4,21*	0,91*	
Iperensione	-	-	-	-	3,72*	-1,79	1,2	17,97*	-4,33*	-1,85	
CRP	-	-	-	-	-	-	-	-	-	-	
Progressione c-IMT	2,84*	0,79	0,06	-	1,01	0,62	-0,73	1,22	0,87	-0,43	
c-IMT	1,02	1,59	5,74*	-	12,77*	8,93*	4,96*	9,47*	5,62*	1,36	
Sgp130	-	-	-	-	-	-	-	-	-	-	
Globuli bianchi	-	-	-	-	-	-	-	-	-	-	
IL6	-	-	-	-	-	-	-	-	-	-	
Infiammazione	-	-	-	-	1,16	-0,78	3,34*	2,15*	10,92*	-7,76*	
Effetti indiretti											
	Diabete	Ipercolesterolemia	Iperensione	Progressione c-IMT	Età	Sesso	Professioni	Lattitudine	Fumo	Attività fisica	
Diabete	-	-	-	-	-	-	-	-	-	-	
Eventi vascolari maggiori	0,6	1,17	1,9	0,03	2,33*	0,62	2,9*	1,59	4,08*	-3,42*	
Ipercolesterolemia	-	-	-	-	-	-	-	-	-	-	
Iperensione	-	-	-	-	-	-	-	-	-	-	
CRP	-	-	-	-	1,16	-0,78	3,31*	2,15*	9,92*	-7,35*	
Progressione c-IMT	1,02	1,58	5,16*	-	8,02*	5,77*	4,82*	5,22*	4,45*	-1,28	
c-IMT	-	-	-	-	2,25*	-1,83	1,99*	3,49*	1,92	-3,32*	
Sgp130	-	-	-	-	0,73	-0,58	0,86	0,84	0,9	-0,89	
Globuli bianchi	-	-	-	-	1,16	-0,78	3,32*	2,14*	9,69*	-7,4*	
IL6	-	-	-	-	1,15	-0,78	3,25*	2,1*	8,77*	-6,84*	
Infiammazione	-	-	-	-	-	-	-	-	-	-	

\* gli effetti statisticamente significativi (p-value <0.05)

## **Discussione**

In questo studio abbiamo mostrato come l'utilizzo di metodologie statistiche innovative possa rappresentare un valore aggiunto rispetto alle metodologie statistiche 'classiche' nell'analisi dei dati epidemiologici.

In particolare, abbiamo illustrato come l'implementazione della Path Analysis e dei SEM mostri un quadro completo ed esaustivo delle relazioni esistenti tra tutti i fattori, non solo di quelle esistenti tra variabili indipendenti (fattori di rischio) e variabili dipendenti di interesse (c-IMT, progressione c-IMT o eventi vascolari maggiori).

Nei Path Diagram mostrati, infatti, si sono valutati il c-IMT e la sua progressione sia come variabili dipendenti (end-point surrogati degli eventi vascolari maggiori), sia che come variabili indipendenti (quali fattori di rischio o cause dirette degli eventi).

Il vantaggio dell'utilizzo della metodologia statistica avanzata, in questo studio, è stato inoltre quello di poter analizzare, in un unico modello, relazioni che, in ambito classico, sono studiate in modelli differenti (regressione lineare, regressione logistica, modelli di Cox).

Infine, ha messo in evidenza anche interessanti relazioni tra i diversi fattori di rischio. In linea di principio ciò sarebbe possibile anche con la metodologia classica, ma si dovrebbero implementare molti modelli di regressione univariata, tanti quanti sono le variabili. Tuttavia, questo non sarebbe ancora sufficiente, in quanto tutti questi modelli di regressione univariata non tengono conto dell'effetto spiegato da tutte le altre variabili. Questo è uno dei vantaggi principali della Path Analysis / SEM rispetto alle regressioni multiple.

## Relazioni Regressioni vs Relazioni stimate con Path Analysis e SEM.

In questo studio, nelle prime analisi, abbiamo implementato diversi modelli di regressione. Nello specifico, quattro regressioni lineari univariate ed una regressione lineare multivariata. Le regressioni lineari univariate sono state impiegate per valutare la relazione tra i singoli biomarcatori di stato di infiammazione (4 Biomarcatori = 4 modelli di regressione lineare univariata) e markers di aterosclerosi clinica, c-IMT, valutato, nel nostro studio, come  $IMT_{mean\_max}$ . abbiamo così individuato una relazione positiva tra Interleuchina-6, globuli bianchi e c-IMT, un'associazione negativa tra sgp130 e c-IMT ed una relazione positiva, ma non significativa tra proteina C-reattiva e c-IMT.

Successivamente, abbiamo implementato un modello di regressione lineare multivariata per valutare l'effetto di ogni singolo biomarcatore di infiammazione, al netto dell'effetto degli altri biomarcatori. Ci siamo qui 'imbattuti' nel paradosso di Simpson (situazione in cui la relazione tra due variabili appare modificata, o perfino invertita, a causa dell'aggiustamento per altre covariate incluse nell'analisi). Infatti, nel modello multivariato, non solo la Proteina C-reattiva diventa significativa, ma il coefficiente cambia di segno e, inaspettatamente, l'associazione con la c-IMT diventa negativa.

Abbiamo quindi calcolato uno score di 'stato infiammatorio' combinando i 4 biomarcatori di infiammazione, ciascuno pesato in base alla sua relazione con il c-IMT basale.

Sempre in ambito di analisi statistica classica, abbiamo calcolato le correlazioni con i fattori di rischio ed implementato delle analisi di incidenza (curva di Kaplan Meier) e modelli di COX. Possiamo quindi affermare che lo score infiammatorio è correlato con c-IMT ed associato agli eventi, indipendentemente dai fattori di rischio considerati (fumo, diabete, ipertensione, ipercolesterolemia, età, sesso e latitudine).

A questo punto, però ci si potrebbe porre una ulteriore domanda: queste associazioni sono dirette o sono il frutto delle relazioni esistenti tra i fattori di rischio, relazioni che non riusciamo a cogliere con le analisi implementate fino ad ora?

Per rispondere a questa domanda abbiamo implementato dei modelli di Path Analysis.

La Figura 14 mostra simultaneamente le relazioni tra:

- score infiammatorio e c-IMT, progressione c-IMT, eventi vascolari maggiori;
- fattori di rischio 'classici' e score infiammatorio;
- fattori di rischio 'classici' e c-IMT, progressione c-IMT, eventi vascolari maggiori;
- c-IMT, progressione c-IMT vs eventi vascolari maggiori;
- Fattori di rischio 'classici' tra di loro

Sicuramente, in ambito classico, non avremmo potuto evidenziare tutte queste relazioni in un unico modello.

Solo per valutare i singoli fattori di rischio tradizionali e lo score infiammatorio, servirebbero 7 modelli di regressione lineare univariata (per le relazioni tra i fattori) più un modello di regressione lineare multivariata, che includa i 7 fattori di rischio 'classici' e lo score come variabile dipendente. quindi, implementando ben 8 modelli otteniamo le informazioni riguardanti solo di una parte delle relazioni visualizzate nel Path diagram.

Il vantaggio di utilizzare una Path analysis per visualizzare simultaneamente tutte le relazioni sopradescritte è evidente.

## Score di stato di infiammazione, c-IMT, progressione c-IMT ed eventi vascolari maggiori.

Anche analizzando lo score di stato infiammatorio, con il Path Diagram si evidenziano molte relazioni interessanti.

La Figura14 mostra la relazione tra score infiammatorio, c-IMT, progressione della c-IMT, eventi vascolari maggiori, fattori di rischio 'classici', variabili socio-economiche e attività fisica.

Il Path Diagram mostra un ruolo centrale dello score di stato di infiammazione: in primis, esso ha un effetto diretto sia sul c-IMT (relazione positiva,  $\beta=2.56$ ) sia sull'end-point combinato eventi vascolari maggiori ( $\beta=9.22$ ).

Inoltre, lo score dello stato di infiammazione è influenzato in modo diretto da età, sesso, latitudine e fumo ( $\beta=2.50$ ,  $\beta=2.31$ ,  $\beta=6.65$ ,  $\beta=6.81$ , rispettivamente).

Da notare che, in questo diagramma, i 'classici' fattori di rischio (Fumo, Diabete, Iperensione, Ipercolesterolemia, Età, Sesso e Latitudine) non hanno un effetto diretto sugli eventi vascolari maggiori, ma molti di loro (Iperensione, Età, Sesso e Latitudine) hanno un effetto indiretto, che si esplica attraverso altre variabili intermedie, ad esempio, l'iperensione ha il suo effetto indiretto sugli eventi, mediato dal c-IMT.

Età, sesso, latitudine e Fumo, invece, influenzano gli eventi vascolari tramite due variabili intermedie (score di stato infiammatorio e c-IMT). Inoltre età e latitudine sembrano influenzare indirettamente gli eventi e il c-IMT, attraverso l'iperensione, sulla quale hanno un effetto diretto ( $\beta=3.18$ ,  $\beta=16.57$ , rispettivamente).

Analizzando la progressione c-IMT, essa non ha effetto diretto sugli eventi vascolari maggiori. Non si osserva inoltre un effetto diretto significativo tra score infiammatorio e progressione c-IMT, ma un solo effetto indiretto mediato dal c-IMT basale ( $\beta$  effetto indiretto=2.50).

Infine, ipertensione, età, sesso, professioni, latitudine e fumo, manifestano un effetto indiretto vs la progressione c-IMT, mediato da c-IMT o/e Diabete.

## **Confronto tra lo score di stato di infiammazione e la variabile latente 'Infiammazione'.**

Infine, abbiamo implementato un modello SEM introducendo una nuova variabile: una variabile latente 'Inflamazione'. come già illustrato, questa variabile è stata costruita ipotizzando delle relazioni teoriche con i 4 biomarcatori di infiammazione.

La variabile latente è stata inserita nel modello di Path Analysis completo comprendente sia il c-IMT che la sua progressione della c-IMT, gli eventi vascolari maggiori, i fattori di rischio 'classici', le variabili socio-economiche, l'attività fisica ed i biomarcatori di stato di infiammazione.

Focalizzandoci sulla variabile latente, possiamo notare che il suo effetto diretto sugli eventi vascolari maggiori è minore rispetto all'effetto diretto visto tra score empirico di stato infiammatorio e tali eventi ( $\beta=4.24$  vs.  $\beta=9.22$ ; il confronto è immediato dato che si tratta di stime standardizzate).

Come lo score di stato infiammatorio la variabile latente è influenzata direttamente da fumo e latitudine ma sono presenti alcune differenze nelle relazioni con attività fisica, lavoro, età e sesso.

Anche la variabile latente, come lo score, non influenza direttamente, ma indirettamente, la progressione c-IMT, attraverso c-IMT.

Da notare che gli eventi vascolari maggiori, sono influenzati direttamente solo dalla variabile latente 'Inflamazione'. Fumo, attività fisica, Lavoro influenzano indirettamente gli eventi, attraverso la variabile latente.

Come già abbiamo anticipato precedentemente, questo tipo di modello non soddisfa tutte e tre i criteri di adattamento (CFI dovrebbe essere  $> 0.9$  ed è uguale a 0.82), per questo motivo non è totalmente un modello affidabile.

In conclusione, confrontando l'utilizzo di una variabile latente calcolata dalle equazioni strutturali con uno score empirico di infiammazione, costruito con criteri simili ad altri score ampiamente utilizzati nella pratica clinica, si può affermare che, dal punto di vista della visualizzazione delle relazioni rilevanti presenti nel modello, i risultati sono

abbastanza simili. Tuttavia, lo score empirico sembra adattarsi meglio ai dati rispetto alla variabile latente. Inoltre, e questo è il punto più importante, quest'ultima è di difficile, se non impossibile, applicazione in clinica, in quanto il suo algoritmo di calcolo non è esplicito, mentre lo score può essere direttamente calcolato in ogni laboratorio partendo dai valori dei biomarcatori.

## Limitazioni

Lo studio condotto in questo progetto e le conclusioni da esso derivate, presentano alcune limitazioni:

- Sia i modelli di regressione multipla 'classici' che la Path analysis e i SEM , comprendono solo alcuni possibili confondenti e sono stati implementati su un numero limitato di variabili. Non si può, quindi, escludere un effetto confondente residuo di altri fattori non misurati. Tuttavia, le analisi riportate sono state condotte principalmente con un scopo illustrativo e modelli più complessi possono certamente essere implementati sia con metodi classici che innovativi.
- Lo score che riflette lo stato infiammatorio dei pazienti è stato valutato su 4 biomarcatori infiammatori, non abbiamo potuto includere altre interleuchine infiammatorie che non erano state rilevate nello studio. Altre variabili di tipo 'infiammatorio' erano presenti nel database, ma con molti dati mancanti.
- Tra i biomarcatori compresi nello score, solo 2 (la Proteina C-reatina e Globuli Bianchi) sono utilizzati nelle analisi di routine.

Tuttavia, la forza e la consistenza dei dati presentati in questo studio, incentivano l'utilizzo di metodologie statistiche innovative, oltre che l'utilizzo di un possibile score di stato di infiammazione per perfezionare la stima del rischio personale verso gli eventi vascolari maggiori.

## **Conclusioni e rilevanza**

I nostri risultati indicano che i metodi statistici innovativi, quali la Path Analysis e SEM, possono essere molto utili nelle ricerche biomediche.

La Path Analysis e i SEM, infatti, ci hanno mostrato tutte le relazioni che eravamo interessati a studiare in un unico grafico e hanno identificato gli effetti indiretti, che non siamo in grado di valutare con i modelli di regressione multipla 'classici'.

Dall'altro lato, le tecniche statistiche innovative nell'ambito della ricerca medica, non devono essere soltanto dei bellissimi modelli, ma deve anche essere possibile applicarli nella pratica clinica.

Si può, quindi, affermare che gli sviluppi della metodologia statistica dovrebbero andare di pari passo con il loro utilizzo nelle ricerche biomediche, ma prestando attenzione alla plausibilità del loro impiego nella pratica clinica.

Infine, i nostri risultati indicano che l'utilizzo di uno score globale di stato infiammatorio può rilevare associazioni con fattori di rischio ed end-point clinici e subclinici con maggiore potenza statistica e capacità predittiva rispetto ai singoli biomarcatori.

Naturalmente, serviranno ulteriori studi per validare tale score e renderlo utilizzabile nella pratica clinica.

## **Ringraziamenti**

Un sentito ringraziamento al Chiar.mo Prof. Giovanni Di Minno, mio tutor negli anni dedicati a questo dottorato di ricerca.

Grazie alla chiar.ma prof.ssa Elena tremoli, direttore scientifico del Centro Cardiologico Monzino, che mi ha dato la possibilità di entrare a fare parte del team di ricerca dell'Istituto.

Grazie di cuore al Dr. Fabrizio Veglia, responsabile dell'unità di Biostatistica, dove in questi anni ho avuto la fortuna di lavorare, formandomi sotto la sua guida e crescendo grazie alla sua fiducia.

Un ringraziamento al Prof. Damiano Baldassarre, coordinatore del gruppo IMPROVE, fonte dei dati utilizzati per lo sviluppo del mio progetto di dottorato.

Grazie alla Dott.sa Bruna Gigante per avermi accolto e supportato durante la mia esperienza presso il Karolinska Institute di Stoccolma-Svezia.

Infine, un ringraziamento speciale ai colleghi dell'unità di Biostatistica per il supporto e la fattiva collaborazione di questi anni trascorsi insieme.

## Referenze

- (1999). "Decline in deaths from heart disease and stroke--United States, 1900-1999." *MMWR Morb Mortal Wkly Rep* 48(30): 649-656.
- (2016). Sewall Wright, path analysis and d-separation. *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference* with R. B. Shipley. Cambridge, Cambridge University Press: 56-86.
- Altman, D. G. (1981). "Statistics and ethics in medical research. VIII-Improving the quality of statistics in medical journals." *Br Med J (Clin Res Ed)* 282(6257): 44-47.
- Altman, D. G. and S. N. Goodman (1994). "Transfer of technology from statistical journals to the biomedical literature. Past trends and future predictions." *JAMA* 272(2): 129-132.
- Amarenco, P., P. C. Lavallee, et al. (2016). "One-Year Risk of Stroke after Transient Ischemic Attack or Minor Stroke." *N Engl J Med* 374(16): 1533-1542.
- Anderson, T. W. (1958). *An introduction to multivariate statistical analysis*. New York,, Wiley.
- Baldassarre, D., K. Nyssonen, et al. (2010). "Cross-sectional analysis of baseline data to identify the major determinants of carotid intima-media thickness in a European population: the IMPROVE study." *Eur Heart J* 31(5): 614-622.
- Baldassarre, D., F. Veglia, et al. (2013). "Progression of carotid intima-media thickness as predictor of vascular events: results from the IMPROVE study." *Arterioscler Thromb Vasc Biol* 33(9): 2273-2279.
- Bernetti, M., R. Abbate, et al. (2011). "[Carotid and femoral intima-media thickness as an early atherosclerotic marker. Advantages and limits]." *G Ital Cardiol (Rome)* 12(1): 72-81.
- Biasucci, L. M., G. Liuzzo, et al. (2000). "Inflammation and acute coronary syndromes." *Herz* 25(2): 108-112.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, Wiley.
- Criqui, M. H. (2001). "Peripheral arterial disease--epidemiological aspects." *Vasc Med* 6(3 Suppl): 3-7.
- Davies, M. G. (2012). "Critical limb ischemia: cell and molecular therapies for limb salvage." *Methodist Debaque Cardiovasc J* 8(4): 20-27.
- De Lorenzo, A., L. Soldati, et al. (2016). "New obesity classification criteria as a tool for bariatric surgery indication." *World J Gastroenterol* 22(2): 681-703.
- Easton, J. D., J. L. Saver, et al. (2009). "Definition and evaluation of transient ischemic attack: a scientific statement for healthcare professionals from the American Heart Association/American Stroke Association Stroke Council; Council on Cardiovascular Surgery and Anesthesia; Council on Cardiovascular Radiology and Intervention; Council on Cardiovascular Nursing; and the Interdisciplinary Council on Peripheral Vascular Disease. The American Academy of Neurology affirms the value of this statement as an educational tool for neurologists." *Stroke* 40(6): 2276-2293.
- Ebrahim, S., O. Papacosta, et al. (1999). "Carotid plaque, intima media thickness, cardiovascular risk factors, and prevalent cardiovascular disease in men and women: the British Regional Heart Study." *Stroke* 30(4): 841-850.

- Edlow, J. A. (2018). "Managing Patients With Transient Ischemic Attack." *Annals of Emergency Medicine* 71(3): 409-415.
- Feigin, V. L., M. H. Forouzanfar, et al. (2014). "Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010." *Lancet* 383(9913): 245-254.
- Giles, M. F. and P. M. Rothwell (2007). "Risk of stroke early after transient ischaemic attack: a systematic review and meta-analysis." *Lancet Neurol* 6(12): 1063-1072.
- Glass, C. K. and J. L. Witztum (2001). "Atherosclerosis. the road ahead." *Cell* 104(4): 503-516.
- Goldberger, A. (1972). "Structural Equation Methods in the Social Sciences." *Econometrica* 40(6): 979-1001.
- Heinrich, P. C., I. Behrmann, et al. (1998). "Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway." *Biochem J* 334 ( Pt 2): 297-314.
- Hirsch, A. T., Z. J. Haskal, et al. (2006). "ACC/AHA 2005 Practice Guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic): a collaborative report from the American Association for Vascular Surgery/Society for Vascular Surgery, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, and the ACC/AHA Task Force on Practice Guidelines (Writing Committee to Develop Guidelines for the Management of Patients With Peripheral Arterial Disease): endorsed by the American Association of Cardiovascular and Pulmonary Rehabilitation; National Heart, Lung, and Blood Institute; Society for Vascular Nursing; TransAtlantic Inter-Society Consensus; and Vascular Disease Foundation." *Circulation* 113(11): e463-654.
- Hocking, R. (1976). "A biometrics invited paper. The analysis and selection of variables in linear regression." *International Biometrics Society* 32: 1-49.
- Hollander, M., M. L. Bots, et al. (2002). "Carotid plaques increase the risk of stroke and subtypes of cerebral infarction in asymptomatic elderly: the Rotterdam study." *Circulation* 105(24): 2872-2877.
- Lambert, M. A. and J. J. Belch (2013). "Medical management of critical limb ischaemia: where do we stand today?" *J Intern Med* 274(4): 295-307.
- Leng, G. C., F. G. Fowkes, et al. (1996). "Use of ankle brachial pressure index to predict cardiovascular events and death: a cohort study." *BMJ* 313(7070): 1440-1444.
- Liuzzo, G., L. M. Biasucci, et al. (1994). "The prognostic value of C-reactive protein and serum amyloid a protein in severe unstable angina." *N Engl J Med* 331(7): 417-424.
- Meaney, C., R. Moinuddin, et al. (2016). "Text mining describes the use of statistical and epidemiological methods in published medical research." *J Clin Epidemiol* 74: 124-132.
- Naka, T., N. Nishimoto, et al. (2002). "The paradigm of IL-6: from basic science to medicine." *Arthritis Res* 4 Suppl 3: S233-242.
- Newman, A. B., L. Shemanski, et al. (1999). "Ankle-arm index as a predictor of cardiovascular disease and mortality in the Cardiovascular Health Study. The Cardiovascular Health Study Group." *Arterioscler Thromb Vasc Biol* 19(3): 538-545.
- Nietert, P. J., A. E. Wahlquist, et al. (2013). "Characteristics of recent biostatistical methods adopted by researchers publishing in general/internal medicine journals." *Stat Med* 32(1): 1-10.

- Norgren, L., W. R. Hiatt, et al. (2007). "Inter-Society Consensus for the Management of Peripheral Arterial Disease (TASC II)." *J Vasc Surg* 45 Suppl S: S5-67.
- Norheim, O. F., P. Jha, et al. (2015). "Avoiding 40% of the premature deaths in each country, 2010-30: review of national mortality trends to help quantify the UN sustainable development goal for health." *Lancet* 385(9964): 239-252.
- O'Connell, J. B. and W. J. Quiñones-Baldrich (2009). "Proper Evaluation and Management of Acute Embolic versus Thrombotic Limb Ischemia." *Seminars in Vascular Surgery* 22(1): 10-16.
- Odden, M. C., P. G. Coxson, et al. (2011). "The impact of the aging population on coronary heart disease in the United States." *Am J Med* 124(9): 827-833 e825.
- Pearson, T. A., G. A. Mensah, et al. (2003). "Markers of inflammation and cardiovascular disease: application to clinical and public health practice: A statement for healthcare professionals from the Centers for Disease Control and Prevention and the American Heart Association." *Circulation* 107(3): 499-511.
- Pedhazur, E. J. (1998). "Multiple Regression in Behavioral Research."
- Pignoli, P., E. Tremoli, et al. (1986). "Intimal plus medial thickness of the arterial wall: a direct measurement with ultrasound imaging." *Circulation* 74(6): 1399-1406.
- Resnick, H. E., R. S. Lindsay, et al. (2004). "Relationship of high and low ankle brachial index to all-cause and cardiovascular disease mortality: the Strong Heart Study." *Circulation* 109(6): 733-739.
- Ridker, P. M. (2003). "Rosuvastatin in the primary prevention of cardiovascular disease among patients with low levels of low-density lipoprotein cholesterol and elevated high-sensitivity C-reactive protein: rationale and design of the JUPITER trial." *Circulation* 108(19): 2292-2297.
- Ridker, P. M., J. E. Buring, et al. (2007). "Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score." *JAMA* 297(6): 611-619.
- Ridker, P. M., M. Cushman, et al. (1997). "Inflammation, aspirin, and the risk of cardiovascular disease in apparently healthy men." *N Engl J Med* 336(14): 973-979.
- Rosamond, W., K. Flegal, et al. (2008). "Heart disease and stroke statistics--2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee." *Circulation* 117(4): e25-146.
- Ross, R. (1999). "Atherosclerosis--an inflammatory disease." *N Engl J Med* 340(2): 115-126.
- Roth, G. A., M. H. Forouzanfar, et al. (2015). "Demographic and epidemiologic drivers of global cardiovascular mortality." *N Engl J Med* 372(14): 1333-1341.
- Roth, G. A., C. Johnson, et al. (2017). "Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015." *J Am Coll Cardiol* 70(1): 1-25.
- Rothwell, P. M., M. F. Giles, et al. (2007). "Effect of urgent treatment of transient ischaemic attack and minor stroke on early recurrent stroke (EXPRESS study): a prospective population-based sequential comparison." *Lancet* 370(9596): 1432-1442.
- Schuett, H., R. Oestreich, et al. (2012). "Transsignaling of interleukin-6 crucially contributes to atherosclerosis in mice." *Arterioscler Thromb Vasc Biol* 32(2): 281-290.
- Simon, A., J. Gariépy, et al. (2002). "Intima-media thickness: a new tool for diagnosis and treatment of cardiovascular risk." *J Hypertens* 20(2): 159-169.

- Sörbom, D. (1975). "DETECTION OF CORRELATED ERRORS IN LONGITUDINAL DATA." *British Journal of Mathematical and Statistical Psychology* 28(2): 138-151.
- Stary, H. C., A. B. Chandler, et al. (1995). "A definition of advanced types of atherosclerotic lesions and a histological classification of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association." *Circulation* 92(5): 1355-1374.
- Stary, H. C., A. B. Chandler, et al. (1994). "A definition of initial, fatty streak, and intermediate lesions of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association." *Circulation* 89(5): 2462-2478.
- Stead, L. G., M. F. Bellolio, et al. (2009). "Evaluation of transient ischemic attack in an emergency department observation unit." *Neurocrit Care* 10(2): 204-208.
- Stein, C. M., N. J. Morris, et al. (2017). "Structural Equation Modeling." *Methods Mol Biol* 1666: 557-580.
- Strasak, A. M., Q. Zaman, et al. (2007). "The Use of Statistics in Medical Research." *The American Statistician* 61(1): 47-55.
- Streiner, D. L. (2005). "Finding our way: an introduction to path analysis." *Can J Psychiatry* 50(2): 115-122.
- Taback, N. and M. K. Krzyzanowska (2008). "A survey of abstracts of high-impact clinical journals indicated most statistical methods presented are summary statistics." *J Clin Epidemiol* 61(3): 277-281.
- van der Meer, I. M., M. L. Bots, et al. (2004). "Predictive value of noninvasive measures of atherosclerosis for incident myocardial infarction: the Rotterdam Study." *Circulation* 109(9): 1089-1094.
- Warr, P. B., R. E. Lee, et al. (1969). "A note on the factorial nature of the F and D scales." *Br J Psychol* 60(1): 119-123.
- Wright, S. (1918). "On the Nature of Size Factors." *Genetics* 3(4): 367-374.
- Wright, S. (1921). "Correlation and causation." *Journal of Agricultural Research, Washington* 20: 557-585.
- Wright, S. (1923). "The Theory of Path Coefficients a Reply to Nilés's Criticism." *Genetics* 8(3): 239-255.
- Yergens, D. W., D. J. Dutton, et al. (2014). "An overview of the statistical methods reported by studies using the Canadian community health survey." *BMC Med Res Methodol* 14: 15.