

IL MODELLO DI REGRESSIONE QUANTILE NELL'ANALISI DELLE DETERMINANTI DELLA QUALITÀ DELLA VITA IN UNA POPOLAZIONE ANZIANA

S. Broccoli, G. Cavrini, M. Zoli

1. INTRODUZIONE

In Italia gli over 65 rappresentano il 20% dell'intera popolazione. Grazie al miglioramento delle procedure diagnostiche e terapeutiche si assiste ad un allungamento della vita media, che non è sempre accompagnato da un equivalente miglioramento della qualità della vita.

Il concetto di “Qualità della Vita” comprende la salute fisica, lo stato psicologico, il livello di autonomia, le relazioni sociali e con l'ambiente dell'individuo (Costituzione OMS, 1948). Questo concetto multidimensionale si trova spesso riassunto nella sigla Health Related Quality of Life (HRQoL).

La sfida del terzo millennio è, quindi, quella di aggiungere qualità di vita agli anni piuttosto che aggiungere anni alla vita (Ricci, 2003). Per perseguire una politica di questo tipo la ricerca epidemiologica dovrà essere rivolta ad individuare i fattori determinanti di una soddisfacente qualità della vita negli anziani, con l'obiettivo di promuoverli nella popolazione in modo mirato e costante.

Il presente lavoro, partendo da dati raccolti su 5256 soggetti ultra sessantacinquenni nell'ambito di uno studio longitudinale denominato “Studio Pianoro”, si propone proprio di indagare quali, tra numerosi fattori, contribuiscono al miglioramento della qualità della vita percepita dall'anziano.

Lo strumento scelto per ottenere una misura di HRQoL è il questionario EQ-5D. Lo strumento è generico, cioè è valido per la popolazione generale. Inoltre, nei lavori di Brazier *et al.* (1996), Holland *et al.* (2004) e Haywood *et al.* (2005) si dimostra che il questionario EQ-5D è uno strumento con buone proprietà nel caso in cui la popolazione di riferimento sia costituita da soggetti anziani.

La metodologia statistica adottata è la regressione quantile, che consente di valutare l'influenza delle determinanti sui diversi quantili della distribuzione condizionata della variabile di *outcome*. In altre parole, la regressione quantile lavora localmente su ogni singolo punto della distribuzione della variabile dipendente. Questa caratteristica è particolarmente interessante negli studi sulla qualità della vita e sullo stato di salute percepito, nei quali gli indicatori di una cattiva salute sono localizzati nella coda della distribuzione.

2. LO “STUDIO PIANORO”

I dati sono stati raccolti nell'ambito di uno studio longitudinale concepito e organizzato dal Dipartimento di Medicina Interna, Cardioangiologia ed Epatologia della Facoltà di Medicina e Chirurgia dell'Università di Bologna, con la collaborazione del Dipartimento di Scienze Statistiche e della Facoltà di Scienze Motorie dell'Università di Bologna, dell'Azienda ASL di Bologna, dei Medici di Medicina Generale di Pianoro e dei Comuni di Pianoro, Sasso Marconi e Zola Predosa. Lo studio, di durata quinquennale, si propone di dimostrare come un incremento dell'attività fisica in una popolazione anziana contribuisca a migliorarne lo stato di salute, con vantaggi in termini di qualità della vita e di costi socio-sanitari. Oltre a questo, si indaga su effetti ancora poco conosciuti dell'allenamento aerobico su una serie di variabili connesse con la sindrome metabolica e l'arteriosclerosi. Obiettivo secondario dello studio è la misura della prevalenza dei tradizionali fattori di rischio cardiovascolare.

I soggetti coinvolti sono gli ultra sessantacinquenni¹ residenti in tre comuni della provincia di Bologna: Pianoro, Sasso Marconi e Zola Predosa. La popolazione trattata, cioè sottoposta ad incentivazione dell'attività fisica, è rappresentata dai residenti nel comune di Pianoro; i residenti negli altri due comuni costituiscono la popolazione di controllo.

Per acquisire conoscenze relative alle caratteristiche socio-demografiche, agli stili di vita e allo stato di salute della popolazione oggetto del nostro studio, ai 9644 residenti nei tre comuni è stato inviato un questionario postale. Inoltre, i soli residenti di Pianoro, sono stati invitati (tramite lettera) a sottoporsi gratuitamente ad una visita medica e ad accertamenti diagnostici, al fine di valutare l'idoneità all'attività fisica, più o meno intensa.

I soggetti risultati “idonei” ai diversi esami clinici sono stati invitati a partecipare a un programma di incentivazione dell'attività fisica. Essi hanno potuto scegliere tra le seguenti attività: passeggiate, percorso vita, bicicletta su circuito ad hoc, nuoto, ballo. Queste attività costituiscono un generico “trattamento” e, al termine dei cinque anni di durata dello studio, si verificherà se l'attività è stata svolta e con quale frequenza. I controlli, cioè i residenti di Sasso Marconi e Zola Predosa, al contrario, non sono stati incentivati in alcun modo e hanno ricevuto il solo questionario postale.

Ad oggi lo studio è in corso ed è attiva la promozione dell'attività fisica tra gli anziani di Pianoro (Cavrini *et al.*, 2005).

Nella prima fase dello studio epidemiologico, di tipo trasversale, il tasso di risposta ai questionari inviati è stato pari al 54,5%. I risultati del presente lavoro sono stati ottenuti analizzando i 5256 questionari restituiti.

Nel questionario, costituito da 8 pagine, sono state richieste informazioni socio-demografiche, antropometriche, stili di vita, quantificazioni dell'attività fisica (PASE), misure di qualità della vita percepita (EQ-5D), attività elementari della vita quotidiana (ADL-IADL), presenza di alcune patologie, numero di cadute e di fratture negli ultimi 3 anni.

¹ Alla data di inizio dello studio, dicembre 2003, quindi nati prima del 31 dicembre 1938.

Per migliorare il tasso di risposta, i questionari sono stati inviati assieme ad una busta pre-affrancata per la restituzione degli stessi, e ad una lettera di accompagnamento, firmata dal responsabile della ricerca, nella quale si sono illustrati gli obiettivi e l'utilità dell'indagine; inoltre si è garantito l'anonimato, ovvero che le informazioni fornite sarebbero state utilizzate esclusivamente a fini scientifici e di programmazione sanitaria.

2.1. Informazioni sociodemografiche, stili di vita e stato di salute

Oltre a sesso, età, stato civile (coniugato, vedovo e una categoria che include celibe/nubile, separato e divorziato) e titolo di studio (classificato in "Basso", se meno di 5 anni di studio, "Medio" per anni di studio compresi tra 5 e 8, "Alto" se più di 8 anni), è stato chiesto al rispondente se vive solo (Sì/No) e se svolge ancora un'attività lavorativa (Sì/No).

Per quanto riguarda gli stili di vita sono state raccolte informazioni relative alle abitudini al fumo (Non fumatore/Ex-fumatore/Fumatore), al consumo di alcool e alla massa corporea: ai fumatori è stato chiesto il numero di sigarette fumate al giorno e da quanti anni si ha questa abitudine; per valutare il consumo di alcol è stato chiesto il numero di bicchieri di vino/superalcolici bevuti in un giorno; infine, con il peso e l'altezza riportati dal rispondente, è stato calcolato il Body Mass Index (BMI²), successivamente trasformato, secondo la categorizzazione dell'Organizzazione Mondiale della Sanità, in: sottopeso (BMI minore di 18,5), normopeso (BMI compreso tra 18,5 e 24,9), sovrappeso (BMI compreso tra 25 e 29,9) e obeso (BMI maggiore o uguale a 30).

Lo stato di salute dei rispondenti è stato valutato considerando la presenza/assenza di diabete, ipertensione, ipercolesterolemia, familiarità di infarto, interventi cardiovascolari (specificando quali), dolori muscolari (specificando le diverse parti del corpo interessate), cadute o fratture negli ultimi tre anni e i farmaci assunti.

Infine, per valutare l'autonomia funzionale sono state inserite le versioni auto-compilate dei questionari Activities of Daily Life (ADL) (Katz e Akpom, 1976) e Instrumental Activities of Daily Life (IADL) (Lawton e Brody, 1969). Con riferimento ai punteggi ottenuti rispetto a queste due scale i soggetti sono stati suddivisi in "non autonomi" (presenza di almeno un problema funzionale) e "autonomi".

2.2. Qualità della vita associata alla salute (HRQoL)

L'EQ-5D è uno strumento generico utilizzato per misurare la qualità della vita percepita.

La prima versione del questionario è stata proposta nel 1990 dal Gruppo Euro-Qol ed era costituita da 6 item; successivamente è stata modificata nel 1991 nella versione definitiva con 5 item (Brooks *et al.*, 1996). Nel corso di questi anni tale strumento è stato ampiamente validato e utilizzato in numerosi studi epidemiologici (Kind *et al.*, 1998; Johnson *et al.*, 1998) e, grazie alla sua facilità di compilazione, è particolarmente appropriato nelle indagini che coinvolgono le persone anziane.

² Il Body Mass Index è dato dal rapporto tra peso (in Kg.) e altezza (in metri) al quadrato.

La qualità della vita viene valutata considerando la presenza di problemi su cinque dimensioni, fondamentali per la HRQoL: mobilità, cura della persona, attività abituali, dolori e/o malessere, ansia e/o depressione. Ogni dimensione viene descritta da tre possibili modalità di risposta: nessun problema, problemi moderati, problemi estremi.

Attraverso l'unione delle risposte viene individuato un "profilo di salute" del soggetto, che permette di rilevare le frequenze dei problemi per dimensione e gravità e di classificare il rispondente in uno dei 243 stati di salute possibili (35). Il profilo di salute può essere ulteriormente elaborato attraverso uno specifico algoritmo che, applicando decrementi per ciascun tipo di problema rilevato, riassume le informazioni raccolte in un punteggio sintetico, denominato EQ-5D Index, e fornisce una misura sintetica dello stato di salute percepito. L'implementazione di questo algoritmo prevede che ad ogni stato di salute sia associato uno specifico peso, calcolato per una popolazione generale utilizzando tecniche proprie delle analisi costo-utilità. Ad oggi non sono ancora stati calcolati i pesi per tutti i Paesi in cui è disponibile l'EQ-5D, compresa l'Italia. In questi casi, solitamente, si utilizzano i pesi anglosassoni.

Il questionario EQ-5D include anche una scala-termometro (Visual Analogue Scale) con cui esprimere la valutazione globale del proprio stato di salute. Il risultato della scala EQ-VAS, che varia tra 0 (il peggior stato di salute possibile) e 100 (il miglior stato di salute possibile), riassume in sé la "qualità della vita" percepita per il profilo di salute definito nella prima parte dello strumento.

Dal momento che nella nostra popolazione l'indice sintetico EQ-5D è stato calcolato utilizzando pesi non italiani, nel presente lavoro si è preferito scegliere come variabile di *outcome* il punteggio EQ-VAS.

2.3. *Physical Activity Scale for the Elderly (PASE) per quantificare l'attività fisica*

Per misurare l'attività fisica svolta si è utilizzato uno strumento denominato PASE (Physical Activity Scale for the Elderly), validato per persone anziane e in grado di rilevare le attività condotte nell'arco dei sette giorni precedenti la compilazione del questionario (Washburn *et al.*, 1993; Dinger *et al.*, 2004).

Strumenti di rilevazione dell'attività fisica adatti per una popolazione generale non sono sufficientemente sensibili ed accurati se utilizzati in una popolazione anziana. Infatti, per quest'ultima, anche semplici gesti quotidiani o brevi passeggiate non possono essere trascurati nel calcolo dell'ammontare dell'esercizio fisico.

Il questionario è composto da due parti: nella prima vengono valutate le attività generiche, considerando separatamente attività domestiche pesanti e leggere, attività di lavoro non sedentario, attività agricole o di giardinaggio e assistenza agli altri; nella seconda parte sono considerate le attività svolte nel tempo libero, con particolare riferimento a passeggiate, jogging, bicicletta, nuoto, ginnastica a casa o in palestra, danza e sport.

Il punteggio PASE Totale si calcola moltiplicando l'ammontare di tempo speso in ogni attività per un peso assegnato ad ogni item e sommando per tutte le attività. L'algoritmo di calcolo prevede che la durata e la frequenza delle attività generi-

che non siano incluse nel calcolo: queste attività vengono registrate solo come eseguite o meno (Washburn *et al.*, 1999; Schuit *et al.*, 1997).

Occorre sottolineare che il questionario PASE non stima l'energia spesa ma produce un punteggio che consente di effettuare confronti.

3. LA REGRESSIONE QUANTILE

Per valutare l'effetto sulla misura della qualità della vita delle numerose covariate a disposizione e la forza di questa relazione è stato utilizzato il modello di regressione quantile.

Questa scelta è motivata anche dalla distribuzione della variabile di *outcome*, il punteggio EQ-VAS. L'istogramma dell'EQ-VAS (Figura 1) ha un andamento "a picchi" (frequenze alte in corrispondenza di alcuni valori e frequenze quasi nulle nell'intervallo tra questi) e non normale; in questo caso è preferibile considerare una regressione definita "robusta" come, ad esempio, la regressione quantile che consente di lavorare localmente su ogni singolo punto della distribuzione. La regressione OLS (Ordinary Least Squares) consentirebbe di modellare una misura di sintesi, la media, ma non terrebbe conto della forma distributiva della variabile di outcome, come è invece possibile fare considerando una regressione quantile. Oltre a questa motivazione strettamente tecnica, l'utilizzo della regressione quantile ha permesso un'analisi più approfondita e dettagliata del fenomeno (Paragrafo 4.2).

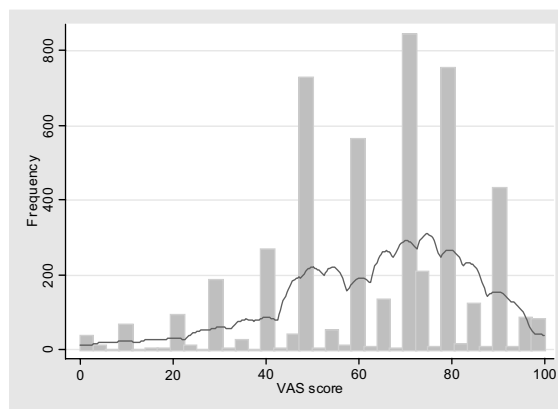


Figura 1 – Istogramma di frequenza e stima Kernel dell'EQ-VAS.

La regressione quantile è stata presentata per la prima volta nel 1978 da Roger Koenker e Gilbert W. Basset, i quali hanno osservato che “*dai metodi classici di regressione, l'unica informazione che si ottiene sulla relazione tra y e il vettore delle covariate X , è il modo in cui la media di Y varia al variare di X* ” (Koenker e Basset, 1978).

Il grande vantaggio della regressione quantile è la possibilità di stimare l'intera distribuzione dei quantili condizionati della variabile risposta, così da poter studiare l'influenza delle variabili esplicative sulla forma della distribuzione di Y . In altre parole, la stima di un valore (media condizionata) viene sostituita dalla stima di 99 valori (quantili condizionati).

Altre vantaggiose caratteristiche della regressione quantile possono essere così sintetizzate (Buchinsky, 1998):

- a) Il problema di minimo da cui si ottengono le stime dei parametri può essere risolto efficacemente con metodi di programmazione lineare (LP).
- b) Essa fornisce stime robuste del vettore dei coefficienti, non sensibili ad outliers nei valori della variabile dipendente.
- c) Quando i termini di errore non sono normalmente distribuiti, gli stimatori forniti dalla regressione quantile possono essere più efficienti degli stimatori dei minimi quadrati.
- d) Se si osservano diverse stime per diversi quantili, si può comprendere come cambia l'influenza delle covariate sulla variabile dipendente, nei vari punti della distribuzione quantile condizionata.
- e) Lo stimatore L, basato su una combinazione lineare di stimatori delle varie regressioni quantili, è sempre più efficiente dello stimatore dei minimi quadrati.

Per la formulazione del modello Koenker e Basset (1978) propongono una procedura alternativa a quella classica per il calcolo dei quantili, basato su un metodo di ottimizzazione. Più precisamente, in analogia a quanto succede per la media campionaria, che può essere definita come la soluzione del problema di minimizzazione della somma di scarti al quadrato, qui possiamo definire ogni singolo quantile come la soluzione del seguente problema di minimo

$$\hat{Q}_Y(\tau) = \underset{\xi_\tau \in \mathbb{R}}{\operatorname{argmin}} \left\{ \sum_{i \in \{i | Y_i \geq \xi_\tau\}} \tau |Y_i - \xi_\tau| + \sum_{i \in \{i | Y_i < \xi_\tau\}} (1-\tau) |Y_i - \xi_\tau| \right\} \quad (1)$$

La formula diventa immediatamente intuitiva se si fa riferimento al caso della mediana, che sappiamo essere quel valore che minimizza la somma degli scarti in valore assoluto. Ponendo $\tau=0,5$ si ottiene infatti

$$\hat{Q}_Y(0,5) = \underset{\xi_{0,5} \in \mathbb{R}}{\operatorname{argmin}} \left\{ \sum_i |Y_i - \xi_{0,5}| \right\} \quad (2)$$

Questa formulazione più complessa per esprimere i quantili risulta particolarmente utile per passare direttamente all'ambito regressivo.

Si consideri il modello di regressione lineare classico

$$y_i = x_i' \beta + u_i \quad \text{con } i=1, \dots, n \quad (3)$$

assunto ($E(u_i | x_i) = 0$), allora la media di y_i condizionata a x_i è data da $(x_i' \beta)$.

Il vettore dei parametri β è stimato con il metodo dei minimi quadrati

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^k}{\operatorname{argmin}} \left\{ \sum_i (y_i - x_i' \beta)^2 \right\} \quad (4)$$

In modo totalmente analogo può essere formulato il modello di regressione quantile:

$$y_i = x_i' \beta_\tau + u_{i\tau} \quad \text{con } i = 1, \dots, n \text{ e } \tau = 1, \dots, 99 \quad (5)$$

assunto ($Q_\tau(u_{i\tau} | x_{i\tau}) = 0$), allora il τ -esimo quantile di y_i condizionato ad x_i è dato da $(x_i' \beta_\tau)$.

Il vettore dei parametri β_τ viene stimato risolvendo un problema di minimo che è espresso come un problema di programmazione lineare (LP) e quindi efficacemente risolto (Buchinsky, 1998).

Per ogni quantile τ il coefficiente stimato indica, per unità di variazione della x a cui è associato, di quanto varia il τ -esimo quantile della y a parità di tutte le rimanenti covariate (Koenker, 2004).

Senza entrare nel dettaglio si ricorda che la stima degli standard error in questo modello regressivo è ancora oggetto di ricerca; sono comunque già presenti in letteratura numerosi algoritmi, tra cui si ricordano i metodi basati sulla stima diretta della matrice di covarianza asintotica, i metodi basati su ricampionamenti bootstrap e il metodo basato sull'inverso del test dei ranghi (Koenker, 1994).

In questo studio la procedura utilizzata per la costruzione del modello è una procedura di tipo *forward*: non sono state considerate le variabili con coefficienti non significativi per i quantili 25°, 50° e 75°. Tutte le variabili sono state centrate rispetto alla propria media, in modo da ottenere un valore della costante che rispecchi il valore del τ -esimo quantile della VAS per un soggetto che presenta caratteristiche medie per ogni variabile (Koenker, 2004). Il metodo di programmazione lineare di stima dei parametri è il metodo di Frisch-Newton, scelto perché risulta essere il più rapido per un campione di osservazioni che vanno da 1000 a 10000. Il metodo utilizzato per il calcolo degli standard error, e quindi dei test di significatività, è il metodo bootstrap di ricampionamento delle coppie X-Y. Le repliche bootstrap sono state fissate a 500.

L'analisi è stata condotta utilizzando la Library "quantreg" del software R-Project 2.0.1 e il software STATA/SE 9.0 (Stata Corporation, Texas, USA).

4. RISULTATI

4.1. *Analisi descrittiva*

I 5256 rispondenti (45,4% maschi) presentano un'età media pari a 74,5 anni (SD=7,1): 73,8 anni i maschi (SD=6,6) e 75,1 anni le femmine (SD=7,4).

Più della metà dei soggetti è coniugata: l'80,3% degli uomini e il 50,5% delle donne. Questa differenza può essere spiegata dalla durata della vita media, più alta per le donne, e dalla consistente presenza di vedove tra queste ultime (43,8% contro il 12,1% dei maschi). Il 73,6% degli anziani intervistati ha un livello di istruzione basso (meno di 5 anni di scuola), e solo il 9,2% dei soggetti svolge ancora un'attività lavorativa (Tavola 1).

TAVOLA 1
*Frequenze assolute e punteggio EQ-VAS per
 caratteristiche demografiche, stili di vita e stato di salute*

<i>Caratteristiche sociodemografiche</i>	N° di soggetti	EQ-VAS Score	
		Media (SD)	Mediana (IR)
Età (Anni)	N (%)		
65 – 69	1566 (29.8)	70.9 (17.2)	70 (60-80)
70 – 74	1353 (25.7)	66.0 (18.3)	70 (50-80)
75 – 79	1080 (20.7)	62.6 (19.7)	70 (50-80)
80 and over	1250 (23.8)	53.7 (22.7)	50 (40-70)
Livello di istruzione			
Basso (<5 anni)	3869 (73.6)	62.0 (20.6)	65 (50-80)
Medio (5-8 anni)	737 (14.0)	67.5 (19.2)	70 (57.5-80)
Alto (>8 anni)	606 (11.5)	72.0 (17.2)	75 (60-85)
<i>Missing Value</i>	44 (0.8)		
Stato civile			
Coniugato	3365 (64.0)	66.1 (10.3)	70 (50-80)
Celibe, separato, divorziato	322 (6.1)	64.3 (19.7)	70 (50-80)
Vedovo	1547 (29.4)	59.0 (21.9)	60 (50-75)
<i>Missing Value</i>	22 (0.4)		
Attività lavorativa			
Sì	484 (9.2)	73.6 (15.5)	80 (65-85)
No	4626 (88.0)	62.8 (20.7)	70 (50-80)
<i>Missing Value</i>	146 (2.8)		
<i>Fattori di rischio</i>	N° di soggetti	EQ-VAS Score	
	N (%)	Media (SD)	Mediana (IR)
Body Mass Index(Kg/m ²)			
Sottopeso (<=18.5)	104 (2.0)	48.9 (28.0)	50 (30-70)
Normopeso(18.5-24.9)	1974 (37.6)	64.0 (21.1)	70 (50-80)
Sovrappeso(25-29.9)	2199 (41.8)	65.3 (19.1)	70 (50-80)
Obesi (>= 30)	696 (13.2)	63.3 (19.3)	65 (50-80)
<i>Missing Value</i>	283 (5.4)		
Fumo			
Mai fumato	2953 (56.2)	63.0 (21.1)	70 (50-80)
Ex- fumatore	1647 (31.3)	65.1 (19.2)	70 (50-80)
Fumatore	499 (9.5)	66.7 (18.9)	70 (55-80)
<i>Missing Value</i>	157 (3.0)		
Consumo di alcool			
Astemio	2193 (41.7)	59.8 (22.0)	60 (50-80)
Moderato consumatore	2746 (52.2)	66.9 (18.6)	70 (52-80)
Forte consumatore	192 (3.7)	70.0 (18.6)	75 (60-80)
<i>Missing Value</i>	125 (2.4)		
PASE			
Inattività	1195 (22.7)	54.2 (25.3)	55 (40-75)
Attività fisica scarsa	1449 (27.6)	60.6 (19.3)	60 (50-75)
Attività fisica moderata	1339 (25.5)	68.0 (16.3)	70 (60-80)
Attività fisica intensa	1273 (24.2)	71.9 (15.7)	70 (60-80)

Segue Tavola 1

Stato di salute	N° di soggetti	EQ-VAS Score	
	N (%)	Media (SD)	Mediana (IR)
ADL			
Indipendenti	3621 (68.9)	69.3 (16.7)	70 (60-80)
Alcuni Problemi	1580 (30.1)	51.4 (22.5)	50 (40-70)
Missing Value	55 (1.0)		
IADL			
Indipendenti	3067 (58.4)	70.7 (15.9)	70 (60-80)
Alcuni Problemi	2158 (41.1)	54.2 (22.1)	50.5 (40-70)
Missing Value	31 (0.6)		
Diabetes			
No	4674 (88.9)	64.7 (20.2)	70 (50-80)
Sì	582 (11.1)	57.8 (21.1)	60 (50-70)
Ipercolesterolemia			
No	3755 (71.4)	64.2 (21.1)	70 (50-80)
Sì	1501 (28.6)	63.2 (18.6)	70 (50-80)
Eventi Cardiovascolari			
No	4532 (86.2)	65.8 (19.6)	70 (50-80)
Sì	724 (13.8)	52.5 (21.8)	50 (40-70)
Ipertensione			
No	2565 (48.8)	65.7 (20.9)	70 (50-80)
Sì	2691 (51.2)	62.2 (19.8)	65 (50-80)
Dolori			
No	350 (6.7)	77.0 (16.0)	80 (70-90)
Sì	4491 (85.4)	62.6 (20.0)	68 (50-80)
Missing Value	415 (7.9)		
Cadute			
No	4220 (80.3)	65.8 (19.5)	70 (50-80)
Sì	1036 (19.7)	56.3 (22.2)	60 (41-70)

L'analisi descrittiva mette in rilievo la presenza di problemi legati all'alimentazione: i soggetti normopeso rappresentano solo il 37,6%; consistente è la percentuale di anziani sovrappeso (41,8%) e obesi (13,2%). Queste percentuali sono sovrappubbili a quelle ottenute in altre indagini nazionali (ISTAT, 2002).

I fumatori costituiscono una quota minima degli anziani (9,5%), ma il 31,3% si è dichiarato ex-fumatore (di questi il 52,5% è costituito da uomini). Infine, ancora più bassa è la percentuale di soggetti che può essere considerata forte consumatore di alcool³ (3,7%).

Un altro problema che dobbiamo sottolineare è l'inattività fisica dei nostri anziani: considerando il punteggio PASE si evidenzia che il 22,7% dei soggetti è completamente inattivo; di questi, il 33,7% è uomo e il 13,7% è donna.

Data la particolare popolazione da noi considerata, non ci meraviglia la presenza di soggetti che riportano problemi di autonomia funzionale (30,1% nell' ADL e 41,1% per l'IADL).

³ Il consumo di alcool è stato considerato moderato se inferiore a 2 bicchieri al giorno per le donne o a 4 bicchieri al giorno per gli uomini, alto se maggiore o uguale a questi limiti.

Infine, la Tavola 1 ci consente di valutare la prevalenza di diabetici (11,1%), di ipercolesterolemici (28,6%) e di ipertesi (51,2%). Il 13,8% dei soggetti dichiara di aver subito un evento cardiovascolare, l'85,4% di avere dolori muscolari e il 19,7% di essere caduto negli ultimi tre anni.

Per quanto riguarda la variabile di *outcome*, il punteggio EQ-VAS, il suo valor medio è pari a 63,9 (SD=20,4), con una evidente differenza tra i due sessi: 61,2 per le donne (SD=20,9) e 67,2 per gli uomini (SD=19,3). Tutte le covariate sono risultate significativamente⁴ associate con l'EQ-VAS. Nella Figura 2 sono riportati i valori del punteggio EQ-VAS per i percentili che saranno successivamente considerati nel modello di regressione quantile: la mediana della distribuzione corrisponde al valore 70. Il 25% dei soggetti attribuisce al suo stato di salute un valore minore o uguale a 50, mentre il 75% di essi ha fornito una valutazione inferiore a 80. L'EQ-VAS diminuisce significativamente all'aumentare dell'età e al diminuire del titolo di studio, e le donne vedove si percepiscono peggio rispetto a quelle coniugate. Il punteggio VAS è, in media, più alto per chi svolge ancora un'attività lavorativa. Gli anziani sottopeso riportano valori estremamente bassi di VAS (M=48,9; SD=28,0). Contrariamente all'atteso, fumatori e forti bevitori riportano valori medi di EQ-VAS più alti degli altri.

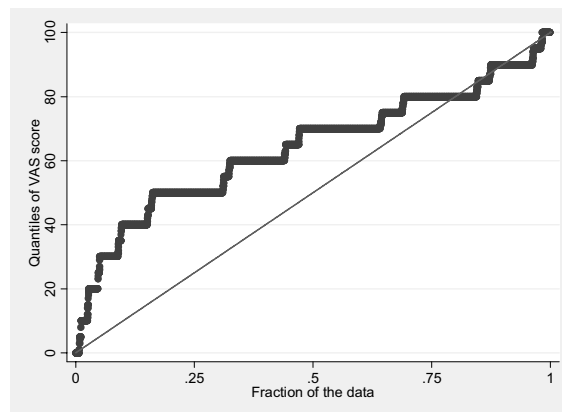


Figura 2 – Rappresentazione dei percentili della variabile EQ-VAS.

Infine, è da sottolineare il fatto che lo stato di salute percepito aumenta all'aumentare dell'attività fisica praticata, diminuisce in presenza di diabete, ipertensione o evento cardiovascolare o in presenza di qualche dolore articolare.

4.2. Risultati del modello di regressione quantile

Nella Tavola 2 sono presentati i risultati ottenuti dalla regressione quantile per il 25°, il 50° e il 75° percentile della distribuzione dell'EQ-VAS Score. In aggiunta, un grafico riassuntivo dell'intero processo di stima delle regressioni quantili (99

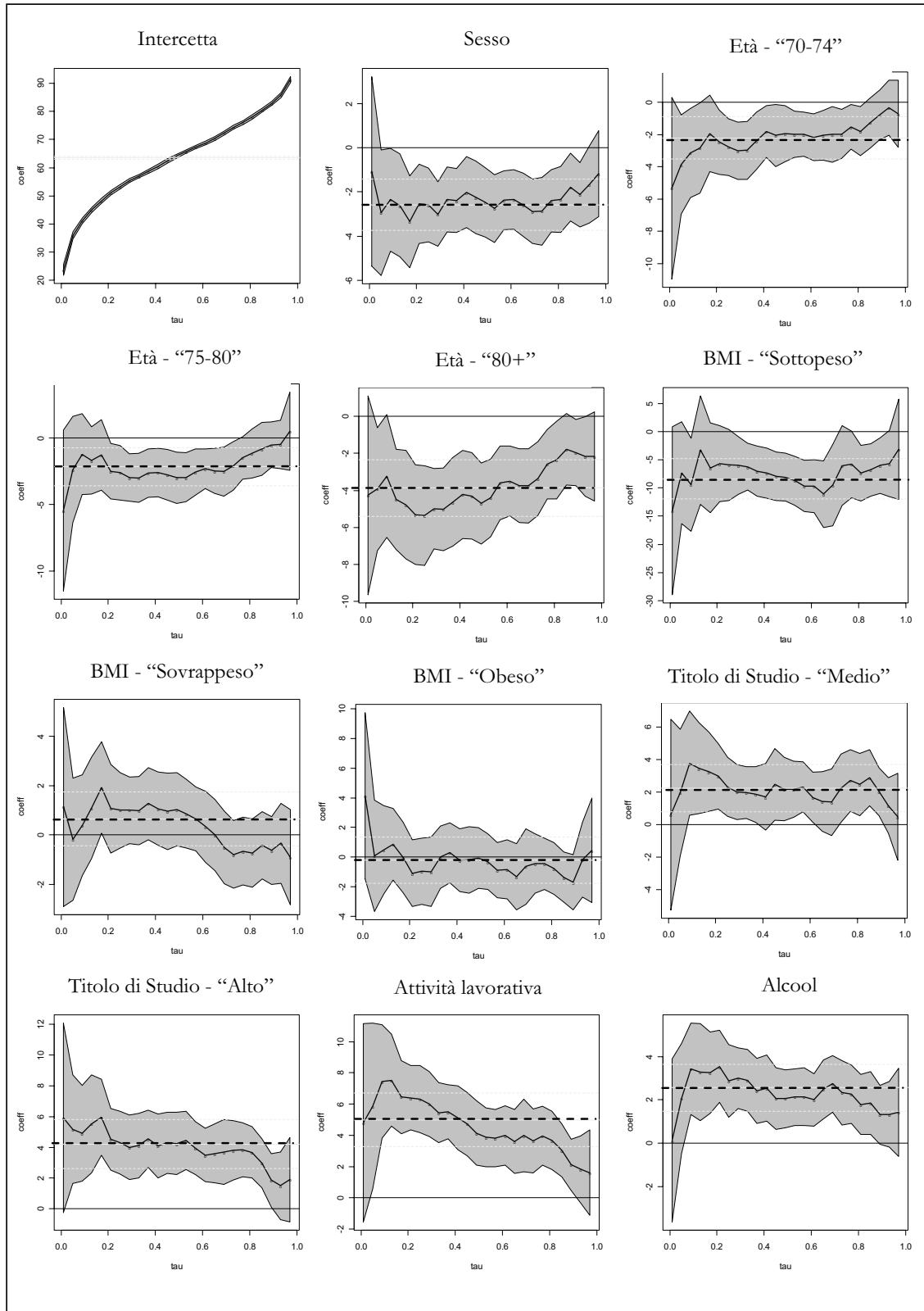
⁴ Si è utilizzata l'ANOVA qualora fosse soddisfatta l'ipotesi di omoschedasticità e il test di Kruskal-Wallis in caso contrario; si è utilizzato il test χ^2 per valutare l'associazione tra due variabili qualitative.

TAVOLA 2

Stima dei coefficienti del modello di regressione quantile per ogni covariata al 25°, 50° e 75° percentile e stima dei coefficienti di un modello di regressione lineare (OLS)

Covariate	OLS	Regressione Quantile					
		q25		q50		q75	
			SE		SE		SE
Sesso (F vs. M)	-2.6 ***	-2.6	*** 0.82	-2.4	*** 0.82	-2.8	*** 0.75
Età							
70-74	-2.2 ***	-2.8	*** 0.91	-2.1	* 0.88	-1.9	* 0.75
75-79	-2.2 ***	-2.6	** 1.00	-3.2	*** 0.96	-1.8	* 0.85
80+	-3.9 ***	-5.4	*** 1.46	-5.0	*** 1.03	-2.9	*** 0.93
BMI (Normopeso)							
Sottopeso	-8.4 ***	-5.9	3.22	-8.3	*** 2.20	-6.2	3.3
Sovrappeso	0.6	1.0	0.75	0.83	0.75	-0.46	0.64
Obeso	-0.2	-0.9	1.20	-0.13	1.03	-0.26	0.91
Titolo di studio (Basso)							
Medio	2.2 **	2.3	* 0.94	2.3	* 0.95	2.5	* 1.05
Alto	4.2 ***	4.3	*** 1.00	4.1	*** 1.00	4.0	*** 0.86
Attività lavorativa (Sì vs No)	5.0 ***	6.3	*** 1.10	4.0	*** 1.04	3.6	*** 0.96
Consumo di alcol (Sì vs No)	2.6 ***	2.9	*** 0.82	2.1	*** 0.67	2.4	*** 0.73
PASE (Inattività)							
Attività fisica scarsa	5.4 ***	5.4	*** 1.09	4.0	*** 1.13	3.0	*** 0.96
Attività fisica moderata	8.7 ***	10.2	*** 1.13	8.1	*** 1.19	5.6	*** 0.91
Attività fisica intensa	10.6 ***	12.1	*** 1.18	9.8	*** 1.07	8.1	*** 0.98
ADL (Sì vs. No)	-8.9 ***	-9.2	*** 0.89	-8.9	*** 0.82	-7.7	*** 0.77
IADL (Sì vs. No)	-7.4 ***	-8.4	*** 0.94	-6.9	*** 0.73	-6.0	*** 0.70
Eventi cardiovascolari (Sì vs. No)	-7.5 ***	-7.8	*** 0.82	-8.6	*** 1.01	-8.1	*** 0.92
Diabete (Sì vs. No)	-3.6 ***	-3.7	*** 1.01	-3.1	* 1.28	-2.6	*** 0.97
Ipertensione (Sì vs. No)	-1.1 **	-0.6	0.74	-1.3	0.73	-1.8	*** 0.62
Colesterolo (Sì vs. No)	-1.0	-0.1	0.75	-1.1	0.73	-1.8	** 0.69
Dolori (Sì vs. No)	-9.2 ***	-10.9	*** 1.12	-10.1	*** 1.07	-7.2	*** 1.00
Cadute (Sì vs. No)	-3.7 ***	-3.9	*** 1.10	-4.4	*** 0.90	-2.4	*** 0.77
Intercetta	63.6 ***	53.2	*** 0.40	64.5	*** 0.33	75.2	*** 0.33

regressioni) è mostrato in Figura 3: i coefficienti stimati per ogni covariata sono riportati in funzione dei quantili (ascisse); le bande grigie ne rappresentano i rispettivi intervalli di confidenza; sono state inoltre riportate una linea orizzontale continua in corrispondenza dello 0 e una linea orizzontale tratteggiata in corrispondenza del valore del coefficiente stimato per quella variabile da un modello di regressione lineare multipla ordinaria. Quest'ultima consente di comprendere e apprezzare maggiormente i risultati ottenuti utilizzando il modello di regressione quantile. Si nota infatti che i risultati di una stima OLS forniscono un solo valore per covariata (linea orizzontale continua), mentre nel caso dei risultati di una regressione quantile i coefficienti stimati si modificano al variare dei quantili della VAS (linea orizzontale spezzata). Da un punto di vista dell'interpretazione, questo ci permette un più preciso commento dell'influenza delle variabili indipendenti sulla variabile di outcome, in particolare per quelle covariate per le quali il valore dei coefficienti stimati cambia significativamente al variare dei quantili (la linea che rappresenta i coefficienti nel grafico non è parallela all'asse delle ascisse, ma pre-



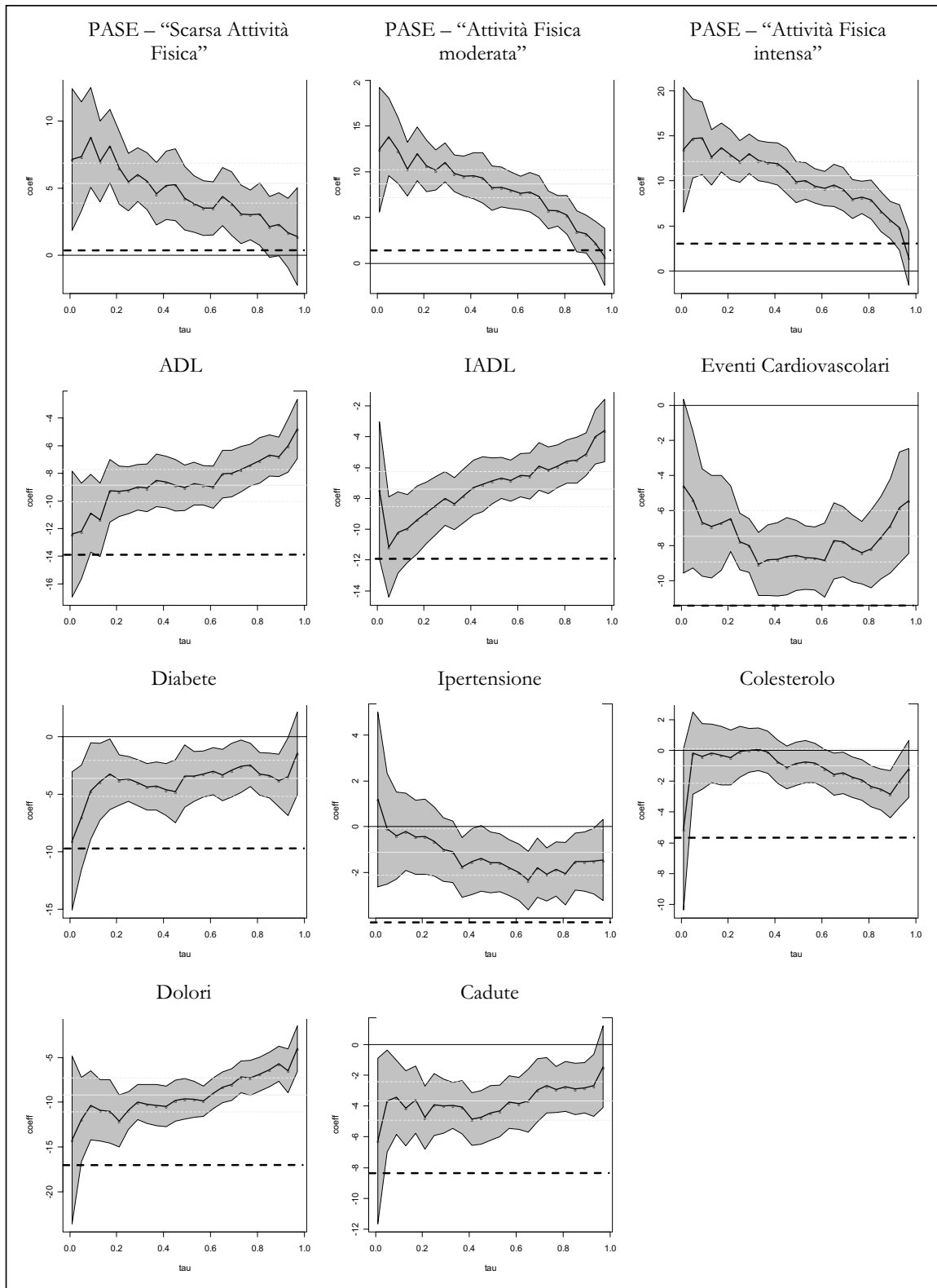


Figura 3 – Per singola covariata coefficienti di 99 regressioni quantili e relativi intervalli di confidenza (bande grigie); coefficiente OLS (linee tratteggiate).

senta una particolare pendenza). Per queste variabili si potrà infatti commentare l'effetto in modo diverso a seconda del livello di qualità della vita riportato dai soggetti.

Il metodo *forward* utilizzato per la scelta delle variabili del modello ha portato ad escludere le variabili "stato civile", "abitudini al fumo" e "vive solo".

L'essere femmina comporta una diminuzione della VAS di circa 2 punti per ogni quantile ($q_{25}=-2,6$; $q_{50}=-2,4$; $q_{75}=-2,8$). Questo conferma quanto già emerso anche in altri studi di popolazione (ISTAT, 2002): a parità delle altre condizioni, le femmine hanno una percezione peggiore della propria salute e della propria qualità della vita.

Naturalmente anche la variabile età gioca un ruolo rilevante nella diminuzione del punteggio VAS: all'aumentare dell'età il punteggio diminuisce (un soggetto con più di 80 anni si attribuisce in mediana un punteggio VAS più basso di 5 punti rispetto ad un soggetto che ha dai 65 ai 69 anni) e il decremento è tanto più intenso quanto più i soggetti si posizionano su quantili bassi ($q_{25}=-5,4$; $q_{50}=-5,0$; $q_{75}=-2,9$).

Sia il possedere un titolo di studio più elevato sia lo svolgere ancora un'attività lavorativa influenzano positivamente la percezione della salute degli anziani.

Anche il consumo di alcool, seppur moderato, contribuisce a migliorare la qualità della vita percepita. Occorre, però, sottolineare che non si conoscono le motivazioni legate al non consumo: in altre parole, non sappiamo se gli astemi siano in questa condizione per scelta o per necessità legate a condizioni critiche di salute.

Per quanto riguarda i problemi legati all'alimentazione, non emergono sostanziali differenze, in termini di punteggio VAS, tra soggetti normopeso e soggetti sovrappeso o obesi. Al contrario, è da sottolineare la differenza significativa tra soggetti normopeso e sottopeso. In questo caso i coefficienti presentano un segno negativo: la qualità della vita è significativamente più bassa per gli anziani sottopeso.

Partendo dai quartili del punteggio PASE è stato possibile suddividere gli anziani in quattro categorie di attività fisica: inattività, attività fisica scarsa, attività fisica moderata e attività fisica intensa (Abete *et al.*, 2001). Dai risultati del modello si evince che anche una attività fisica scarsa, rispetto alla completa inattività, è associata positivamente alla percezione dello stato di salute. Osservando nel grafico (Figura 3 - PASE) l'andamento decrescente dei coefficienti rispetto ai quantili si nota che l'intensità di questa relazione è maggiore per i soggetti che si posizionano in percentili bassi di punteggio VAS ($q_{25}= 5,4$; $q_{50}=4,0$; $q_{75}=3,0$). Inoltre i coefficienti crescono all'aumentare della classe di attività fisica.

Infine, come ci si può aspettare, la mancanza di autonomia, la presenza di patologie, quali eventi cardiovascolari, ipertensione, diabete, ipercolesterolemia, dolori articolari e cadute riducono significativamente i valori della VAS.

5. CONCLUSIONI

Il presente lavoro si colloca nell'ambito dello "Studio Pianoro", il cui principale obiettivo è valutare l'effetto dell'attività fisica sullo stato di salute e sulla qualità della vita percepita dalla popolazione anziana. Attraverso l'analisi dei primi dati raccolti in tale studio ci si propone qui di valutare le relazioni esistenti tra determinanti della salute e qualità della vita. La metodologia statistica utilizzata è il modello di regressione quantile, che consente di stimare la distribuzione dei quantili condizionati della variabile di outcome, ovvero di studiare l'influenza delle caratteristiche socio-demografiche, degli stili di vita e delle patologie sulla forma distributiva del punteggio EQ-VAS.

Tra i risultati più rilevanti che sono stati ottenuti, occorre sottolineare che la presenza di patologie e di dolore contribuiscono a peggiorare la percezione della salute.

Un importante fenomeno indagato che incide positivamente sulla percezione della salute è l'attività fisica. Anche solo una leggera attività fisica aumenta significativamente la percezione della salute per tutti i percentili della distribuzione VAS. Grazie allo strumento statistico utilizzato è stato inoltre possibile capire che questa positiva associazione tra l'attività fisica e la qualità della vita è tanto più intensa quanto più i soggetti si percepiscono peggio. Questi risultati sono decisamente incoraggianti nell'ottica dello studio Pianoro, che si propone di dimostrare proprio che l'attività fisica è fonte di benessere.

In conclusione, questo studio ha mostrato che titolo di studio, attività lavorativa e una sana attività fisica, anche se modesta, associate ovviamente ad un miglioramento dello stato di salute fisica, rappresentano le determinanti più importanti per una percezione migliore della propria qualità della vita.

Dal punto di vista statistico, occorre sottolineare che benchè la regressione quantile sia ampiamente conosciuta ed utilizzata nella letteratura economica ed econometrica, non è stata ancora sufficientemente utilizzata in epidemiologia. In tale ambito, spesso variabili di outcome con distribuzioni fortemente asimmetriche, come l'EQ-VAS, sono comunque analizzate con modelli di regressione OLS. In questo lavoro abbiamo utilizzato la regressione quantile con il duplice obiettivo di valutare l'influenza delle determinanti sui diversi quantili della distribuzione condizionale della VAS e di mostrare la possibilità di utilizzare questi modelli anche in ambito epidemiologico.

Dipartimento di Scienze Statistiche "Paolo Fortunati"
Università di Bologna

SERENA BROCCOLI

Dipartimento di Scienze Statistiche "Paolo Fortunati"
Università di Bologna

GIULIA CAVRINI

Dipartimento di Medicina Interna, Cardioangiologia, Epatologia
Università di Bologna

MARCO ZOLI

RINGRAZIAMENTI

Lo studio è stato finanziato dalla Fondazione Cassa di Risparmio in Bologna e dalla Regione Emilia Romagna, Piani per la Salute

Gruppo Studio Pianoro

Dipartimento di Medicina Interna, Cardioangiologia, Epatologia – University of Bologna
Comitato Direttivo: Marco Zoli (Chief), Giampaolo Bianchi, Donatella Magalotti, Antonio Muscari
Staff Medico: Annalisa Berzigotti, Nicola Castaldini, Grazia Maria Costa, Susanna Dapporto, Gianmarco Drago, Claudia Giannoni, Cosimo Ludovico, Fausta Montesi, Francesco Nicolino, Stefano Ramilli, Dario Sbanò, Paola Zappoli
Staff Tecnico: Raffaella Chianese, Franca Ferri, Giorgia Passerini

Dipartimento di Istologia, Embriologia e Biologia Applicata – University of Bologna
Staff Attività Fisica: Pasqualino Maietta, Claudio Tentoni, Erika Nerozzi, Carlo Ravaoli, Paola Masi, Irene Negrini

Dipartimento di Scienze Statistiche – University of Bologna
Staff Statistico: Giulia Cavrini, Andrea Mattivi, Serena Broccoli

Azienda USL Bologna

Staff Epidemiologico: Marilina Colombo, Paolo Pandolfi, Barbara Pacelli

Staff di Laboratorio: Amedeo Ligabue, Loretta Dini, Annalisa Zacchini

Medici di Medicina Generale: Alberto Melli, John Sonnino, Antonio Ajuti, Stefano Bacci, Luca Bruni, Gemma Chiarini, Giampaolo Corradini, Fulvia Nucci, Elisabetta Ongaro, Elisabetta Scandellari, Serena Selva

Comune di Pianoro

Simonetta Saliera (Sindaco), Gianalberto Cavazza, Antonella Grazia, Maria Pia Mezzini, Daniela Mignogna, Emanuela Torchi

RIFERIMENTI BIBLIOGRAFICI

- P. ABETE N. FERRARA, F. CACCIATORE, E. SAGNELLI, M. MANZI, V. CARNOVALE, C. CALABRESE, D. DE SANTIS, G. TESTA, G. LONGOBARDI, C. NAPOLI, F. RENGO (2001), *High level of Physical activity preserves the cardioprotective effect of preinfarction angina in elderly patients*, "Journal of the American College of Cardiology", 38, art 5, pp. 1357-1365
- R. BROOKS (1996), *EuroQol: The current state of play*, "Health Policy", 37, pp. 53-72
- R. BROOKS, R. RABIN, F. DE CHARRO (2003), *The measurement and valuation of health status using EQ-5D: a European perspective*, Kluwer Academic Publisher
- M. BUCHINSKY (1998), *Recent advances in quantile regression model: a practical guideline for empirical research*, "The Journal of human resources", 33, pp. 88-126
- G. CAVRINI, B. PACELLI, A. MATTIVI, G. BIANCHI, P. PANDOLFI, M. ZOLI (2005), *Benefits of physical activities on the perceived health in elderly A presentation of the study*. "Atti del 7° Convegno Europeo EuroQol", Oslo
- M.K. DINGER, R.F. OMAN, E.L. TAYLOR, S.K.J. ABLE (2004), *Stability and convergent validity of the Physical Activity Scale for the Elderly (PASE)*, "Journal of Sports, Medicine and Physical Fitness", 44, pp. 186-192
- K.L. HAYWOOD, A.M. GARRATT, R. FITZPATRICK (2005), *Quality of life in older people: a structured review of generic self-assessed health instruments*, "Quality of life research", 14, part 7, pp. 1651-1658

- R. HOLLAND, R.D. SMITH, I. HARVEY, L. SWIFT, E. LENAGHAN (2004), *Assessing quality of life in the elderly: a direct comparison of the EQ-5D and AQL*, "Health Economics", 13, part 8, pp. 793-805
- ISTAT (2002), *Le condizioni di salute della popolazione*, in Indagine Multiscopo sulle famiglie "condizioni di salute e ricorso ai servizi sanitari" – Anni 1999-2000
- J.A. JOHNSON, S.J. COONS, A. ERGO, G. SZAVA-KOVATS (1998), *Valuation of EuroQoL (EQ-5D) health states in an adult US sample*, "Pharmacoeconomics", 13, pp. 421-433
- S. KATZ, C.A. AKPOM (1976), *A measure of primary sociobiological functions*, "International Journal of Health Services", 6, pp. 493-508
- P. KIND, P. DOLAN, C. GOUDIX, A. WILLIAMS (1998), *Variations in population health status: results from a United Kingdom national questionnaire survey*, "British Medical Journal", 316, pp. 736-741
- R. KOENKER, G. BASSET (1978), *Quantile regression*, "Econometria", 46, part 1, pp. 33-50
- R. KOENKER (1994), *Confidence interval for regression quantile* in Mandl and Huskova (eds.), "Proceedings of the fifth Prague symposium on Asymptotic statistics", Physica, pp. 349-359
- R. KOENKER (2005), *Quantile regression*, Cambridge University Press, Cambridge
- M.P. LAWTON, E.M. BRODY (1969), *Assessment of older people: self-maintaining and instrumental activities of daily living*, "Gerontologist", 9, pp. 179-186
- P. RICCI (2003): *Salute e qualità della vita: quali servizi per fornire risposte adeguate ai bisogni individuali*, "Atti del Convegno Governare il sistema salute", Roma
- A.J. SCHUIT., E. SCOUTEN, K.R. WESTERTERP, W.H. SARIS (1997), *Validity of the Physical activity Scale for the Elderly (PASE): According to energy expenditure asses by the doubly labelled water method*, "Journal of Clinical Epidemiology", 50, pp. 541-546
- R.A. WASHBURN, K.W. SMITH, A.M. JETTE, C.A. JANNEY (1993), *The physical activity scale for the elderly (PASE): Development and evaluation*, "Journal of Clinical Epidemiology", 46, pp. 153-162
- R.A. WASHBURN, E. MCAULEY, J. KATULA, S.L. MIHALKO, R.A. BOILEAU (1999), *The physical activity scale for the elderly (PASE): Evidence of validity*, "Journal of clinical epidemiology", 52, pp. 643-651

RIASSUNTO

Il modello di regressione quantile nell'analisi delle determinanti della qualità della vita in una popolazione anziana

Obiettivi. L'obiettivo di questo studio è spiegare l'associazione tra alcune importanti covariate e la qualità della vita percepita (HRQoL) in soggetti anziani. **Metodi.** I dati sono stati raccolti nell'ambito di uno studio longitudinale che coinvolge 5256 soggetti ultrasessantacinquenni. La scala-termometro del questionario EQ-5D, la EQ-VAS, è stata utilizzata per ottenere una misura sintetica di qualità della vita. L'EQ-VAS Score è stato modellato utilizzando la regressione quantile. Questo approccio metodologico è stato preferito alla regressione lineare a causa della particolare distribuzione del punteggio EQ-VAS. Le principali covariate inserite nel modello sono: quantità settimanale di attività fisica, capacità di svolgimento delle attività quotidiane (questionari ADL e IADL), presenza di patologie cardiovascolari, diabete, ipertensione e colesterolo, dolori articolare e informazioni socio-demografiche. **Principali Risultati.** 1) Anche valori bassi di attività fisica sono associati significativamente ed in modo positivo alla qualità della vita percepita. 2) Problemi nella attività quotidiane, almeno un evento cardiovascolare e dolori articolari contribuiscono fortemente a ridurre il punteggio EQ-VAS.

SUMMARY

A quantile regression approach to the analysis of the quality of life determinants in the elderly

Objective. The aim of this study is to explain the effect of important covariates on the health-related quality of life (HRQoL) in elderly subjects. *Methods.* Data were collected within a longitudinal study that involves 5256 subjects, aged ≥ 65 . The Visual Analogue Scale included in the EQ-5D Questionnaire, the EQ-VAS, was used to obtain a synthetic measure of quality of life. To model EQ-VAS Score a quantile regression analysis was employed. This methodological approach was preferred to an OLS regression because of the EQ-VAS Score typical distribution. The main covariates are: amount of weekly physical activity, reported problems in Activity of Daily Living, presence of cardiovascular diseases, diabetes, hypercholesterolemia, hypertension, joint pains, as well as socio-demographic information. *Main Results.* 1) Even a low level of physical activity significantly influences quality of life in a positive way; 2) ADL problems, at least one cardiovascular disease and joint pain strongly decrease the quality of life.