RESEARCH PAPER

Evidence on the Bidirectional Relationship Between Health and Life Satisfaction in Older Adults

Darío Moreno-Agostino[1,2,3] · **Francisco José Abad[4]** · **Francisco Félix Caballero[5,6]**

Abstract

Previous research on health and life satisfaction in older adults has suggested a bidirectional relationship. However, most evidence either is based on cross-sectional data, being unsuitable for inferring any directionality on the results, or disregards the within-person stability of both variables over time, thus providing potentially biased results. We analysed data from 11,667 older adults interviewed between 2008 and 2016 within the English Longitudinal Study of Ageing. A health measure including self-reported and measured tests on cognitive and physical performance was computed using a Bayesian multilevel item-response theory approach. Life satisfaction was assessed with the Satisfaction with Life Scale. Cross-lagged models using a Structural Equation Modelling approach were used to analyse the longitudinal interrelation between health and life satisfaction. Models accounting and not accounting for the within-person stability in the measures were computed. Additional sensitivity models were ran using an exclusively self-reported measure of health. Health and life satisfaction were significant longitudinal predictors of one another only when the within-person stability was not considered. The effect of life satisfaction on health became negligible once that stability was modelled. The impact of not accounting for within-person stability was larger when health was measured exclusively with self-reported information. Our study suggests that health and life satisfaction are not bidirectionally related in older adults, but only health predicts a better life satisfaction over time. Thus, policies aimed at fostering older adults' life satisfaction by focusing on health enhancement and maintenance may be fruitful, whereas the opposite may not.

Keywords Subjective wellbeing · Longitudinal study · Satisfaction with life · Functional ability · Healthy ageing

1 Introduction

The global population is ageing. The previsions by the United Nations are that, by 2050, the population aged 60 years or older will reach 2.1 billion (United Nations, 2017). This shift in global demographics is a success story that comes with a major challenge: to ensure

✉ Darío Moreno-Agostino
   d.moreno@ucl.ac.uk

Extended author information available on the last page of the article

🙋 Springer

that those added years are lived with health so older people are able to do what they value (Beard et al., 2016). This idea is captured and extended by the World Health Organization (WHO) in the *World Report on Ageing and Health* (World Health Organization, 2015), where "happiness, satisfaction and fulfilment", also known as subjective wellbeing (SWB) in the literature (Diener, 2006), are identified as the very end of the healthy ageing process. This is consistent with the advice by entities such as the European Commission (2016) or the Organisation for Economic Co-operation and Development (OECD, 2013) to foster subjective wellbeing at the population level.

Many studies have explored the relationship between health and subjective wellbeing in the general population. A meta-analysis of 29 studies showed an overall statistically significant positive relationship between health and subjective wellbeing (Ngamaba et al., 2017). As in most of the SWB literature, the studies included in the meta-analysis namely considered evaluative measures of SWB (e.g. global assessments of the life satisfaction as a whole or general statements about affect, including happiness), which are widely present in population surveys (Dolan & Metcalfe, 2012).

This relationship also takes place in the older population. Even if the relationship between health and SWB is not among the aims of the study, researchers focusing on the determinants of SWB in older adults usually include some measure of health (typically self-reported health, number of chronic conditions, or functional ability) in order to adjust the results for the potential confounding effect of health on those relationships (some recent examples are Lam & García-Román, 2019; Moreno-Agostino et al., 2019, 2020b). Moreover, a plethora of studies have specifically analysed the relationship between health and SWB in older adults, mostly focusing on SWB as the outcome. In line with the abovementioned evidence, these studies have also shown that health predicts a better SWB in older adults from low-, middle-, and high-income countries (Gildner et al., 2019; Lu et al., 2020; Moreno-Agostino et al., 2021a; Puvill et al., 2016, 2019). These findings go in line with a "bottom–up" perspective of SWB, where pleasurable and unpleasurable moments and experiences are what determine the SWB assessments people make (Brief et al., 1993; Diener et al., 2018; Feist et al., 1995); and with what has been called the "disability hypothesis", which posits that health problems can exert a negative effect on SWB (Watson & Pennebaker, 1989). However, the reverse hypothesis (which could be denoted as the "psychosomatic hypothesis") has also been tested, showing that SWB predicts better health in older adults (Miret et al., 2017; Steptoe & Fancourt, 2019). Nevertheless, a common limitation of all these studies is that, due to the use of cross-sectional designs, there is no way to ascertain any directionality in the results, leaving unanswered the question of whether health predicts SWB or vice versa in older adults.

Fewer studies have used a longitudinal approach that allows for exploring a specific directionality in the relationship between health and SWB in older adults. For instance, using a linear mixed modelling approach, Calderón-Larrañaga et al. (2019) recently found that SWB predicted a lower speed of disability accumulation and multimorbidity over time, replicating the same results when taking into account only older adults with no disability or multimorbidity at baseline. On the other hand, studies focusing on the trajectories of SWB in older adults have found that a better health is consistently associated with the best trajectories over time (Lim et al., 2017; Moreno-Agostino et al., 2021b). A common limitation of these longitudinal approaches is that they fail to provide evidence that accounts for the potential bidirectionality in the relationship between SWB and health.

A notable exception to this is the study by Gana et al. (2013), where this bidirectionality is explicitly modelled by using a cross-lagged model in a sample of older adults over five time points. In this article, the authors provide evidence on the highly autoregressive

nature of health and SWB measures (i.e. the impact of the previous state on the next one), along with the relative impact of each of them on the other, showing that poorer health levels are related to subsequent lower levels of SWB, whereas the reverse relationship was found to be statistically non-significant. Nevertheless, previous research has shown that classic cross-lagged models fail to separate the within-person process, which are intended to model, from stable between-person differences (Hamaker et al., 2015). This limitation is especially relevant in this context, since both SWB and health measures have been found to have a relatively high personal stability, also in old age (Daskalopoulou et al., 2019; de la Fuente et al., 2018; Diener et al., 2009; Moreno-Agostino et al., 2020a, 2021b). When this within-person stability is not specifically modelled, that potential stability is then captured by the autoregressive coefficients of the cross-lagged model, which inflates their magnitude and can affect the cross-lagged estimates (Zyphur et al., 2019). Moreover, failing to separate within- and between-person variability does not rule out the possibility that the relationship between time-specific assessments of health and SWB may be explained by broader "top down" individual predispositions to assess more positively or negatively different aspects of their lives (Brief et al., 1993; Diener et al., 2018; Feist et al., 1995), especially if, as is usually the case, health assessments rely exclusively on self-reported information.

Considering the abovementioned literature, there is a need for gathering new evidence on the potentially bidirectional relationship of health and SWB in older adults that overcomes the limitations of previous studies and provides useful insights that can be used to prioritise interventions aimed at fostering the SWB and healthy ageing of the population (European Commission, 2016; World Health Organization, 2017).

Measures of life satisfaction are widely available in population-based studies (Dolan & Metcalfe, 2012), including ageing cohort studies such as the English Longitudinal Study of Ageing (ELSA) (Steptoe et al., 2013). Although SWB has additional components (e.g., experiential and eudemonic wellbeing), the lack of availability of adequate measures of these components in population-based studies makes evaluative wellbeing measures, as life satisfaction, the most suitable choice.

Therefore, the main aim of the present study is to explore the relationship between health and life satisfaction in older adults, accounting for the potential bidirectionality in these effects as well as the within-person stability in both constructs. Since most of the studies show that health is one of the most important predictors of life satisfaction, and the size of the reverse effect is usually of a lower magnitude, the hypothesis is that the effect of health on SWB will be higher than the effect of SWB on health in older adults.

2 Materials and Methods

2.1 Sample and Procedure

Data were drawn from the English Longitudinal Study of Ageing (ELSA), a biannual longitudinal study of a nationally representative sample of the English population aged 50 and older (Steptoe et al., 2013). Waves 4 (2008) to 8 (2016) were considered for this work due to the uninterrupted inclusion of the variables of interest. The initial sample comprised 13,243 older adults. Among those, only participants aged 50+ at baseline, with at least one observation in one of the variables whose longitudinal interrelation was being analysed, and with complete information in the covariates used (i.e., gender and baseline age

and wealth) were included in the corresponding analyses. All participants provided written informed consent. The ELSA study was approved by the National Research Ethics Service (MREC/01/2/91). Additional information on the study and how to access the data can be found on the webpage https://www.elsa-project.ac.uk/.

2.2 Measures

2.2.1 Life Satisfaction

Life satisfaction was assessed by means of the Satisfaction with Life Scale (SWLS) (Diener et al., 1985). In the SWLS, respondents are asked to express their agreement with five statements about their life satisfaction in a 7-point Likert scale from 1 ("Strongly disagree") to 7 ("Strongly agree"). The resulting score ranges from 5 (minimum life satisfaction) to 35 (maximum life satisfaction). Previous research has shown that the SWLS items have an adequate internal consistency (with Cronbach's α values ranging between .80 and .90), and that the scores present a moderate stability over time (even over a decade) as well as reasonable sensitivity to changes over time (Pavot & Diener, 2008). The internal consistency of the SWLS in the five waves considered measured with Cronbach's alpha ranged from .89 (wave 4) to .90 (waves 5, 6, 7, and 8).

2.2.2 Health

Health was operationalised using the analytical approach proposed by Caballero et al. (2017). A set of 45 health-related items covering different health domains [i.e. walking, sight, hearing, balance, dizziness, memory, orientation in time, cognition, pain, energy, sleep, incontinence, mobility, and limitations in Activities of Daily Living (ADLs) and Instrumental Activities of Daily Living (IADLs)], consistently with the WHO recommendations for health measurement (Salomon et al., 2003), were selected. These items included both self-reported questions on functioning and difficulties, and measured tests on cognitive (e.g. verbal fluency or immediate recall) and physical performance (i.e. speed of walk). The items were subsequently used to compute a health score at each time point for each participant in that wave based on a polytomous Bayesian multilevel Item Response Theory (IRT) approach (Fox & Glas, 2001). Previous research has shown that this score is able to predict mortality and institutionalisation better than other frequent operationalisations of health such as the count of chronic conditions (Caballero et al., 2017; de la Fuente et al., 2018). Although not all 45 items were present across all waves, the IRT approach used allowed overcoming this obstacle. The full list of items is included in Table S1 (Supplementary Material). Following the procedure described in Caballero et al. (2017), walking speed was recoded into three groups based on the quartiles of the distribution at each wave (lowest quartile, two central quartiles, and highest quartile), whereas cognitive tasks' scores were recoded into three groups based on their mean and standard deviation (SD) at each wave: high (> one SD above the mean), medium (± one SD around the mean), and low (< one SD below the mean).

2.2.3 Other Variables

Gender and baseline age and wealth were included as covariates affecting health and life satisfaction at all time points in the cross-lagged models. The measure of wealth in

quintiles elaborated by the Institute for Fiscal Studies (2011) for the ELSA study was included in quintiles of net total wealth. This measure reflects the savings, investments, physical wealth and housing wealth after financial debt and mortgage has been subtracted. In those cases where information on wealth was missing in all the waves under study, that information was retrieved if available from previous waves of the ELSA study.

Additionally, a question on self-reported health ("Would you say your health is excellent, very good, good, fair, or poor?"), with scores ranging from 1 ("excellent") to 5 ("poor") was considered for the sensitivity analyses. Response category scores were reversed so higher scores reflected a better self-reported health.

2.3 Statistical Analysis

First, descriptive statistics of the age, gender, and wealth of the analytic and excluded samples in the main analyses were obtained. Significant differences between both samples were analysed by unpaired t-test for continuous variables and chi-square test for categorical ones, using Cohen's d and Cramer's V, respectively, as effect size estimates.

A Bayesian multilevel IRT approach was used in order to obtain the health scores of each individual at each time point in which they participated. A set of four IRT models with increasing complexity were computed in order to select the model with best fit to the data. The resulting expected a posteriori (EAP) was extracted from the model with the best fit (lowest Deviance Information Criterion, DIC). Details on the approach can be find in the Appendix S1 (Supplementary Material).

Although the raw sum scores of the SWLS were used in the subsequent models, we assessed the correlation between these and the factor scores estimated from a confirmatory factor analysis (CFA) model with an adequate level of longitudinal invariance as a way to assess the equivalence between these scores in terms of the rank-ordering of individuals. A model with metric (or weak) longitudinal invariance was deemed sufficient considering that the focus of the study is on the relationships between variables rather than on the mean scores over time (Mulder & Hamaker, 2020). Details on the approach can be found in the Appendix S2 (Supplementary Material).

In order to analyse the longitudinal interrelationship between health and life satisfaction, we computed a set of cross-lagged models. These models are dynamic models that allow assessing the effects of two or more variables on each other over time (Hamaker et al., 2015). We used a structural equation modelling (SEM) approach to specify two different models. The first one, a standard cross-lagged model, was computed namely for comparison purposes with the second model and with previous research (Gana et al., 2013). This first model included both autoregressive effects (the effect of health or life satisfaction at the previous wave $t - 1$ on the same variable at the wave t) and cross-lagged effects (the effect of health or life satisfaction at $t - 1$ on the other variable at t). The covariance between the residual variances of health and life satisfaction at each wave was freely estimated in order to allow for potential within-occasion effects on the observations. The main advantage of the standard specification is its simplicity and transparency, whereas its main downside is that it fails to separate the within-person and between-person variance, even if the multiple observations are nested within individuals, so the within-person stability is absorbed by the autoregressive component of the models, which biases the effect estimates (Hamaker et al., 2015).

A second cross-lagged model was computed in order to overcome this limitation. In addition to the previous model's specifications, this model included a set of additional

latent variables to model additional phenomena. Following the suggestions in Zyphur et al. (2019), we specified two latent variables (η_{health} and $\eta_{life\ satisfaction}$), each of them identified by their respective observed variables with unrestricted factor loadings except for those corresponding to $t = T$ (total number of waves), that were fixed to 1 for identification purposes. These two latent variables had unrestricted covariance among them. By using this strategy, "fixed-effects" due to between-person variability in the repeated measures are captured by these latent variables, thus not biasing the resulting autoregressive and cross-lagged effects. The freely estimated covariance across η_{health} and $\eta_{life\ satisfaction}$ may be especially important when, as it often happens in the literature (and it is specifically modelled in the sensitivity analyses described below), health is operationalised by means of a self-reported item, whose relationship with life satisfaction measures obtained through similar psychological processes may bias the results.

As a sensitivity analysis, two additional models were computed with the same specifications as the two above-mentioned cross-lagged models, using the self-reported measure on health instead of the IRT score.

All cross-lagged models were computed with a SEM approach, using MLR estimator, which provides standard errors that are robust to non-normality (Maydeu-Olivares, 2017). Gender and baseline age and wealth were included as exogenous variables with effects on all health and life satisfaction observations at all time points. Model fit was assessed by means of Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA) values. Following Hu and Bentler (1999) recommendations, CFI and TLI values higher than .95, and RMSEA values lower than .06 were considered to indicate good model fit. In both cases, cross-lagged effects of health and life satisfaction were compared in the standardised solution to ascertain the relative impact of each variable on the other.

Missing data were assumed to be missing at random (MAR), meaning that missingness is ignorable upon conditioning on the observed variables. Full Information Maximum Likelihood (FIML) estimation was used, which allows each case to contribute to the estimation of the parameters for which there are complete data, thus using all available data for parameter estimation and providing unbiased estimates under the MAR assumption (Enders & Bandalos, 2001). Nevertheless, the FIML estimation does not account for missingness in the exogenous variables (i.e., gender and baseline age and wealth). Therefore, considering the substantial subset of participants whose wealth level could not be ascertained, and in order to analyse the potential impact of this missingness in the results, we computed both cross-lagged models (with and without the within-person stability factors) excluding wealth.

Data management was performed with Stata SE version 14 (StataCorp, 2015). Bayesian multilevel IRT and CFA were performed in R version 3.6.3 (R Core Team, 2020), using *sirt* R package version 3.9-4 (Robitzsch, 2020) for IRT and *lavaan* R package version 0.6-7 (Rosseel, 2012) for CFA. Cross-lagged models were performed with Mplus version 7 (Muthén and Muthén 1998–2017). The code is available upon reasonable request.

3 Results

3.1 Sample Characteristics

Among the 13,243 participants aged 50+that comprised the initial sample, 13,233 (99.92%) provided information on more than 50% of the selected health indicators on at least one occasion; 11,779 (88.95%) completed the SWLS questionnaire on at least one occasion; 12,634 (95.40%) provided information on their self-reported health on at least one occasion; and 11,673 (88.14%) had complete information in the covariates (the wealth of the remaining 1570 participants could not be ascertained). The analytic sample of the main models (using the IRT-based health measure and adjusting for gender and baseline age and wealth) comprised 11,667 participants. Table 1 summarises the baseline characteristics of the analytic and excluded samples. Almost all cases in the excluded sample ($n = 1570$, 99.62%) were excluded because their wealth level could not be ascertained. Significant differences were found in the age of the samples, with the included sample being 5.38 years younger ($p < .001$) than the excluded sample. A higher percentage of female was found in the included sample (55.65%) compared to the excluded sample (44.35%). The distributions of the health and life satisfaction scores across the different waves are shown in the Table S2 (Supplementary Material).

3.2 Obtention of the Health Score

The DIC values and the EAP reliability estimates for the different Bayesian multi-level IRT models performed are detailed in Table 2. The best fitting model was the model including

Table 1 Baseline characteristics of the analytic and excluded samples

	Analytic sample (N=11,667)	Excluded sample (N=1576)	t/χ^2	Effect size
Age, M (SD)	65.10 (10.11)	59.72 (8.21)	− 20.25***	− .54
Gender, N (%)				
Female	6456 (55.34)	699 (44.35)	67.43***	.07
Male	5211 (44.66)	877 (55.65)		
Wealth, N (%)				
1st quintile	2209 (18.93)	1 (.06)	3.69	.02
2nd quintile	2373 (20.34)	1 (.06)		
3rd quintile	2281 (19.55)	0 (.00)		
4th quintile	2365 (20.27)	1 (.06)		
5th quintile	2439 (20.91)	3 (.19)		
Missing	0 (.00)	1570 (99.62)	–	–

The analytic sample corresponds to the sample used in the main models with the item response theory-based health score and adjusted for gender and baseline age and wealth

Effect size estimate for age correspond to Cohen's *d*; the remaining effect size estimates correspond to Cramer's *V*, and have been computed with non-missing data only

M mean, *SD* standard deviation, *N* number of cases, *t* unpaired *t*-test, used for continuous variables; χ^2 chi-square test, used for categorical variables

***$p < .001$

Table 2 Information criteria and reliability of the Bayesian multilevel item response theory models of health

Model	DIC	EAP
1: Item-wise item difficulties (b_i)	829,041	.929
2: Model 1 + item-wise item slopes (a_i)	749,890	.952
3: Model 2 + joint standard deviation of item difficulties (σ_b)	609,335	.951
4: Model 2 + item-specific standard deviation of item difficulties (σ_{bi})	546,015	.947

7000 iterations were run, and the first 100 iterations were discarded (burn-in)

DIC deviance information criterion, *EAP* expected a posteriori reliability

item-wise difficulties and slopes, with item-specific standard deviation of item difficulties (model 4), showing the lowest DIC value (546,015) and an adequate EAP reliability (.947). Therefore, this model was use for obtaining the health score of participants.

3.3 Rank-Ordering Equivalence of SWLS Scores

The goodness-of-fit indices of the CFA models computed to assess the longitudinal invariance of the SWLS are shown in Table S3 (Supplementary Material). The metric invariance model showed to adequately reproduce the data, and the correlations found between the predicted factor scores from this model and the raw sum scores used in the subsequent models ranged from .981 to .982, showing an almost identical rank-ordering of participants.

3.4 Cross-lagged Models with the IRT-Based Health Measure

In order to address the aims of the present study, we used the health scores from the selected IRT model along with the SWLS scores to perform the cross-lagged models. The goodness-of-fit indices associated to the cross-lagged models computed with the IRT-based health score are included in the left section of Table 3. The standard cross-lagged model did not show an optimal fit to the data, with some of its fit indices being very close to the habitual cut-off criteria for unacceptable fit. The visual depiction of this cross-lagged model is shown in Fig. 1, including the standardised regression coefficients (β) and the residual correlations at each time point. All autoregressive and cross-lagged effects were found to be positive and significant. The largest effect was found for the autoregressive effect of health, with β estimates ranging from .849 to .871, followed by the autoregressive effect of life satisfaction, which ranged from .694 to .717. Regarding the cross-lagged effects, health effect on life satisfaction (ranging from .099 to .119) was found to be larger than the effect of life satisfaction on health (ranging from .018 to .020).

The cross-lagged model accounting for the within-person stability showed substantially better fit. All goodness-of-fit indices showed an excellent fit of the model to the data, except for the χ^2 estimate (to be expected due to the large sample size). The visual depiction of this model along with the standardised coefficients is shown in Fig. 2. The inclusion of the latent factors capturing the within-person stability in health and life satisfaction over time substantially reduced the autoregressive effects of both variables, which kept being the largest effects in the model. Nevertheless, the cross-lagged

Table 3 Goodness-of-fit indices of the cross-lagged models

	IRT-based health score		Self-reported health score		IRT-based health score (not adjusted for wealth)	
	Standard cross-lagged model	Cross-lagged model with within-person stability factors	Standard cross-lagged model	Cross-lagged model with within-person stability factors	Standard cross-lagged model	Cross-lagged model with within-person stability factors
n	11,667		11,322		13,233	
χ^2 (df)	2100*** (36)	173*** (25)	3659*** (36)	127*** (25)	2288*** (36)	171*** (25)
No. of parameters	59	70	59	70	49	60
CFI	.955	.997	.908	.997	.951	.997
TLI	.906	.990	.808	.992	.911	.992
RMSEA (90% CI)	.070 (.068, .073)	.023 (.019, .026)	.094 (.092, .097)	.019 (.016, .022)	.069 (.066, .071)	.021 (.018, .024)
BIC	538,214	534,975	320,802	316,624	595,299	591,646
SABIC	538,027	534,752	320,615	316,402	595,143	591,456

BIC Bayesian information criterion, CFI comparative fit index, CI confidence interval, df degrees of freedom, SABIC sample size-adjusted Bayesian information criterion, TLI Tucker-Lewis index

***$p < .001$
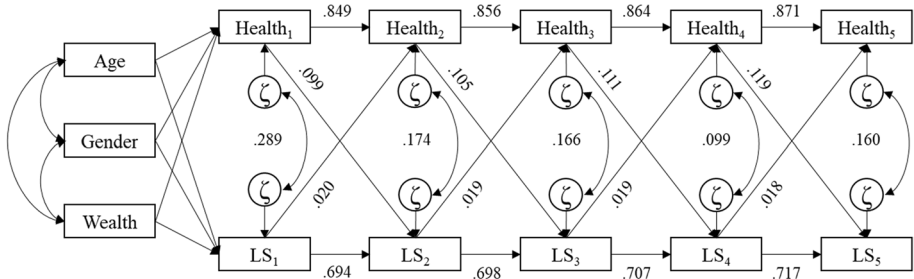
Fig. 1 Standard cross-lagged model between life satisfaction and the IRT-based health measure. *Note*: ζ residuals, *LS* life satisfaction. All coefficients are standardised. The effects of the exogenous variables (i.e. age, gender, and wealth) on health and life satisfaction on later time points (Health$_2$ and LS$_2$ and onwards) are not depicted for the sake of clarity
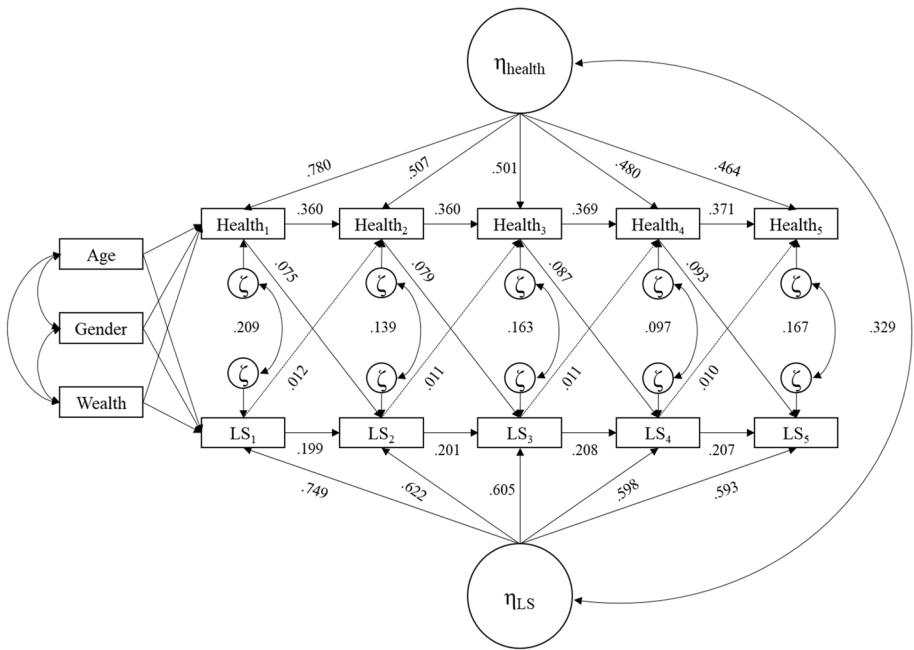

Fig. 2 Within-person stability cross-lagged model between life satisfaction and the IRT-based health measure. *Note*: η within-person stability factors, ζ residuals, *LS* life satisfaction. All coefficients are standardised. The effects of the exogenous variables (i.e. age, gender, and wealth) on health and life satisfaction on later time points (Health$_2$ and LS$_2$ and onwards) are not depicted for the sake of clarity. The dotted lines represent effects with a significance $.01 < p < .05$

effects of life satisfaction on health ranged from $\beta = .010$ to $\beta = .012$ ($p = .019$), whereas its counterpart, the effect of health on life satisfaction, ranged from $\beta = .075$ to $\beta = .093$ ($p < .001$) and was around 7–8 times larger than the former. The two within-person stability latent factors had a positive correlation ($\varphi = .329$).

3.5 Sensitivity Analyses: Cross-lagged Models with Self-reported Health

In order to ascertain to what extent the relationships found would be affected by the use of an exclusively self-reported measure of health instead of the IRT-based measure including both self-reported and measured tests of cognitive and physical functioning, we computed an additional set of models. The goodness-of-fit indices of these sensitivity models are shown in the middle section of Table 3. In this case, the standard cross-lagged model showed a poor fit to the data, whereas the model including the within-person stability latent factors showed an excellent fit. The visual depiction of the standard cross-lagged model computed with the self-reported measure of health is shown in the Figure S1 (Supplementary Material). First, it is important to note that the lack of fit of this model hinders its substantive interpretation. Nevertheless, compared to its counterpart with the IRT-based health score, this model showed a lower autoregressive effect of health (ranging from .629 to .639), thus being lower than that of life satisfaction. Moreover, the cross-lagged effects were in this case very similar across the two variables: the effect of life satisfaction on health ranged from .081 to .084, whereas the effect of health on life satisfaction ranged from .091 to .093, only slightly larger.

The cross-lagged model including the within-person stability latent factors is depicted in Figure S2 (Supplementary Material). Compared to the analogous model with the IRT-based health score, and as in the case of the standard cross-lagged model with the self-reported health measure, the autoregressive effect of health was smaller (ranging from .142 to .145). Regarding the cross-lagged effects, the effect of life satisfaction ($\beta = .012$) turned to be non-significant ($p = .093$) once the within-person stability was accounted for. Notably, the cross-lagged effect of health on life satisfaction was substantially smaller (ranging from .036 to .038) than in the models with the IRT-based measure. The correlation between the two latent factors was larger in this model ($\varphi = .410$).

3.6 Sensitivity Analyses: Cross-lagged Models Without Adjusting for Wealth

To assess the effect of excluding cases with no information on wealth, and in light of the differences found in the comparison between included and excluded samples, an additional set of cross-lagged models were computed without adjusting for that variable. This increased the sample size in 1566 additional cases. The goodness-of-fit indices of these models are shown in the right section of Table 3, and their visual depiction, along with the coefficients, are shown in the Figure S3 (standard cross-lagged model) and Figure S4 (cross-lagged model with within-person stability factor) of the Supplementary Material. The results were very similar to those adjusted for wealth, again showing that, once the within-person stability in the measures was modelled, the cross-lagged effects of life satisfaction on health became very small and, as in the case with the self-reported health measure, not statistically significant ($p = .064$).

4 Discussion

4.1 Main Findings and Interpretation

The present study shows that, while health predicts a higher life satisfaction in older adults, once the within-person stability in both variables is accounted for, the effect of life satisfaction on health becomes negligible. These results were replicated using two different operationalisations of health: a more comprehensive approach including both self-reported and performance information and a single self-reported question on general health. Models not accounting for that stability may find biased results, suggesting that life satisfaction is a relevant predictor of health, especially when using self-reported measures of health, where the results may even suggest that the ability of each variable to predict the other one is equivalent. Altogether, our study reflects the importance of accounting for the within-person stability when analysing the relationship between health and life satisfaction in older adults, most especially when the operationalisation of health relies exclusively on self-reported information.

In this latter case, failing to account for that stability and the relationship between both stability factors can become a greater source of bias, since both constructs are more likely to be related due to unspecific factors such as the response styles (Angelini et al., 2012) or broader "top down" individual predispositions at assessing the personal levels of health and life satisfaction (Diener et al., 2018), given that both variables are exclusively self-reported. In our study, it is likely that such response styles may explain why health and life satisfaction stability factors are more highly related in the models using exclusively self-reported information for operationalising health.

Our results are in line with a previous study by Gana et al. (2013), which used a similar analytical approach and found that health predicts life satisfaction but life satisfaction does not predict health in older adults. Our study provides further evidence on these relationships by specifically accounting for the within-person stability of both measures and using a health operationalisation that includes objective performance information, thus reducing the risk of bias in the estimates. Moreover, our results underpin the relative personal stability of both constructs over time in the older population (Daskalopoulou et al., 2019; de la Fuente et al., 2018; Diener et al., 2009; Moreno-Agostino et al., 2020a, 2021b).

On the contrary, the present study suggests that caution should be taken when interpreting evidence analysing the relationship between health and life satisfaction (Ngamaba et al., 2017), especially evidence reporting that SWB predicts a better health in older adults (Miret et al., 2017; Steptoe & Fancourt, 2019). It is important to note that most of the previous evidence (Miret et al., 2017; Ngamaba et al., 2017) has been conducted with cross-sectional data, and therefore no directionality in the results can be inferred. But even when longitudinal information is used (Steptoe & Fancourt, 2019), the above-mentioned stability in the constructs involved is usually disregarded. By using a robust approach that accounts for the complex existing dynamics across health and life satisfaction in older adults, our study overcomes these limitations, providing a more accurate picture of the underlying relationships.

4.2 Strengths and Limitations

To our knowledge, this is the first study to analyse the relationship between health and life satisfaction in older adults using a methodological approach that accounts for the within-person stability in both constructs, while adjusting the results for potentially relevant confounders such as age, gender, and wealth. Importantly, unlike most of the available literature on this relationship, our approach allows us to model the directionality of the effects in a longitudinal framework, thus leaving less room for alternative explanations. Furthermore, we have used a robust approach to the measurement of health in old age (Caballero et al., 2017; Daskalopoulou et al., 2019), thus avoiding relying exclusively on self-reported information. Nevertheless, by using an alternative approach to the operationalisation of health based exclusively on self-reported information, we have been able to show how greater bias may take place when failing to account for the within-person stability in this habitual case in the literature.

Notwithstanding, the present study should be interpreted on the light of several limitations. First, it is crucial to note that the sample analysed in this study belongs to a population from a high-income Anglo-Saxon country. This may importantly affect the generalisability of the results found, that may have substantially differed had the study been conducted on the older population from a country that is not Western, Educated, Industrialised, Rich, and Democratic (WEIRD), where living conditions are expected to change the way in which health and SWB interrelate over time (Henrich et al., 2010). Moreover, it has been suggested that people from Anglo-Saxon countries may display differences in how their SWB is expected to change over time (Steptoe et al., 2015), although previous research has found similar results when studying older participants from France (Gana et al., 2013). Therefore, replication studies in non-WEIRD older populations are needed in order to generalise the findings of this work. Another limitation of the present study is the exclusive consideration of life satisfaction measures. SWB comprises additional components that, despite being related to life satisfaction, are not strictly equivalent to it (Dolan et al., 2017). The existence of a positive relationship between health and these different components has been proposed in the literature (Moreno-Agostino et al., 2021b; Ngamaba et al., 2017; Steptoe et al., 2015), but the magnitude of the relationships found in the present study may have varied if alternative components of SWB had been considered. Unfortunately, and as it is usual in population-based studies (Dolan & Metcalfe, 2012), only life satisfaction was consistently included and measured over time. Future research may focus on the relationship across health and SWB over time using different components such as experiential or eudemonic wellbeing. It is also important to note that measurement error was not accounted for in the relationships found. This is because, although the higher complexity of the models including the within-person trait-like stability factors allows separating within- and between-person variability (Hamaker et al., 2015; Mulder & Hamaker, 2020), it also makes it difficult to simultaneously estimate the measurement and structural models, being prone to present convergence problems (Usami, 2020; Usami et al., 2019). Nevertheless, methods for estimating longitudinal models on reciprocal relationships are being developed and refined, and future research may replicate these findings embedding measurement models for the variables under study in the structural model testing the reciprocal effects. Finally, the direction of the differences found between the analytic and excluded sample suggests that the latter comprised more men and of younger age, which may limit the generalisability of our results. However, the additional

set of analysis performed without adjusting for wealth, and thus avoiding the exclusion of the great majority of those cases, showed that the results were very similar to those of the main analysis, supporting their robustness.

5 Conclusions

Health and life satisfaction are not bidirectionally related in older adulthood when the within-person stability in both constructs is accounted for, but only health predicts a better life satisfaction over time. Policies aimed at fostering the SWB of the older population may achieve that goal by focusing on health enhancement and maintenance, but, on the contrary, policies aimed at promoting healthy ageing by increasing SWB may not be as effective. Research aimed at analysing health and life satisfaction in older adults, and possibly in the general population, should account for the relative stability of those constructs in order to adequately ascertain the underlying relationships across them and with other variables.

Declarations

Conflict of interest The authors declared that they have no conflict of interest to declare.

References

Angelini, V., Cavapozzi, D., Corazzini, L., & Paccagnella, O. (2012). Age, health and life satisfaction among older Europeans. *Social Indicators Research, 105*(2), 293–308. https://doi.org/10.1007/s11205-011-9882-x

Beard, J. R., Officer, A. M., & Cassels, A. K. (2016). The world report on ageing and health. *The Gerontologist, 56*(Suppl 2), S163–S166. https://doi.org/10.1093/geront/gnw037

Brief, A. P., Butcher, A. H., George, J. M., & Link, K. E. (1993). Integrating bottom–up and top–down theories of subjective well-being: The case of health. *Journal of Personality and Social Psychology, 64*(4), 646–653. https://doi.org/10.1037//0022-3514.64.4.646

Caballero, F. F., Soulis, G., Engchuan, W., Sanchez-Niubo, A., Arndt, H., Ayuso-Mateos, J. L., et al. (2017). Advanced analytical methodologies for measuring healthy ageing and its determinants, using factor analysis and machine learning techniques: The ATHLOS project. *Science and Reports, 7*, 43955. https://doi.org/10.1038/srep43955

Calderón-Larrañaga, A., Vetrano, D. L., Welmer, A. K., Grande, G., Fratiglioni, L., & Dekhtyar, S. (2019). Psychological correlates of multimorbidity and disability accumulation in older adults. *Age and Ageing, 48*(6), 789–796. https://doi.org/10.1093/ageing/afz117

Daskalopoulou, C., Koukounari, A., Wu, Y. T., Terrera, G. M., de la Fuente, J., Caballero, F. F., et al. (2019). Healthy ageing trajectories and lifestyle behaviour: The Mexican health and aging study. *Science and Reports, 9*(1), 11041. https://doi.org/10.1038/s41598-019-47238-w

de la Fuente, J., Caballero, F. F., Sanchez-Niubo, A., Panagiotakos, D. B., Prina, M. A., Arndt, H., et al. (2018). Determinants of health trajectories in England and the US: An approach to identify different patterns of healthy aging. *Journals of Gerontology. Series a, Biological Sciences and Medical Sciences*. https://doi.org/10.1093/gerona/gly006

Diener, E. (2006). Guidelines for national indicators of subjective well-being and ill-being. *Journal of Happiness Studies, 7*(4), 397–404. https://doi.org/10.1007/s10902-006-9000-y

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75.

Diener, E., Lucas, R. E., & Oishi, S. (2018). Advances and open questions in the science of subjective well-being. *Collabra: Psychology*. https://doi.org/10.1525/collabra.115

Diener, E., Lucas, R. E., & Scollon, C. N. (2009). Beyond the hedonic treadmill: Revising the adaptation theory of well-being. In *The science of well-being* (pp. 103–118). Springer.

Dolan, P., Kudrna, L., & Stone, A. (2017). The measure matters: An investigation of evaluative and experience-based measures of wellbeing in time use data. *Social Indicators Research, 134*(1), 57–73. https://doi.org/10.1007/s11205-016-1429-8

Dolan, P., & Metcalfe, R. (2012). Measuring subjective wellbeing: Recommendations on measures for use by national governments. *Journal of Social Policy, 41*(2), 409–427. https://doi.org/10.1017/S0047279411000833

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

European Commission. (2016). *European framework for action on mental health and wellbeing*. EU Joint Action on Mental Health and Wellbeing.

Feist, G. J., Bodner, T. E., Jacobs, J. F., Miles, M., & Tan, V. (1995). Integrating top–down and bottom–up structural models of subjective well-being: A longitudinal investigation. *Journal of Personality and Social Psychology, 68*(1), 138–150. https://doi.org/10.1037/0022-3514.68.1.138

Fox, J.-P., & Glas, C. A. W. (2001). Bayesian estimation of a multilevel IRT model using Gibbs sampling. *Psychometrika, 66*(2), 271–288. https://doi.org/10.1007/BF02294839

Gana, K., Bailly, N., Saada, Y., Joulain, M., Trouillet, R., Herve, C., et al. (2013). Relationship between life satisfaction and physical health in older adults: A longitudinal test of cross-lagged and simultaneous effects. *Health Psychology*, *32*(8), 896–904. https://doi.org/10.1037/a0031656

Gildner, T. E., Snodgrass, J. J., Evans, C., & Kowal, P. (2019). Associations between physical function and subjective well-being in older adults from low- and middle-income countries: Results from the Study on Global AGEing and Adult Health (SAGE). *Journal of Aging and Physical Activity, 27*(2), 213–221. https://doi.org/10.1123/japa.2016-0359

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102–116. https://doi.org/10.1037/a0038889

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X discussion 83–135.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Institute for Fiscal Studies. (2011). *Financial derived variables ELSA waves 1–8*. Retrieved May 25, 2020 from https://www.ifs.org.uk/publications/5749

Lam, J., & García-Román, J. (2019). Solitary day, solitary activities, and associations with well-being among older adults. *The Journals of Gerontology: Series B*. https://doi.org/10.1093/geronb/gbz036

Lim, H. J., Min, D. K., Thorpe, L., & Lee, C. H. (2017). Trajectories of life satisfaction and their predictors among Korean older adults. *BMC Geriatrics, 17*(1), 89. https://doi.org/10.1186/s12877-017-0485-5

Lu, N., Spencer, M., Sun, Q., & Lou, V. W. Q. (2020). Family social capital and life satisfaction among older adults living alone in urban China: The moderating role of functional health. *Aging & Mental Health*. https://doi.org/10.1080/13607863.2019.1709155

Maydeu-Olivares, A. (2017). Maximum likelihood estimation of structural equation models for continuous data: Standard errors and goodness of fit. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(3), 383–394. https://doi.org/10.1080/10705511.2016.1269606

Miret, M., Caballero, F. F., Olaya, B., Koskinen, S., Naidoo, N., Tobiasz-Adamczyk, B., et al. (2017). Association of experienced and evaluative well-being with health in nine countries with different income levels: A cross-sectional study. *Global Health, 13*(1), 65. https://doi.org/10.1186/s12992-017-0290-0

Moreno-Agostino, D., Caballero, F. F., Martin-Maria, N., Tyrovolas, S., Lopez-Garcia, P., Rodriguez-Artalejo, F., et al. (2019). Mediterranean diet and wellbeing: Evidence from a nationwide survey. *Psychology & Health, 34*(3), 321–335. https://doi.org/10.1080/08870446.2018.1525492

Moreno-Agostino, D., Daskalopoulou, C., Wu, Y. T., Koukounari, A., Haro, J. M., Tyrovolas, S., et al. (2020a). The impact of physical activity on healthy ageing trajectories: Evidence from eight cohort studies. *International Journal of Behavioral Nutrition and Physical Activity, 17*(1), 92. https://doi.org/10.1186/s12966-020-00995-8

Moreno-Agostino, D., de la Fuente, J., Leonardi, M., Koskinen, S., Tobiasz-Adamczyk, B., Sanchez-Niubo, A., et al. (2021a). Mediators of the socioeconomic status and life satisfaction relationship in older adults: A multi-country structural equation modeling approach. *Aging & Mental Health, 25*(3), 585–592. https://doi.org/10.1080/13607863.2019.1698513

Moreno-Agostino, D., de la Torre-Luque, A., de la Fuente, J., Lara, E., Martín-María, N., Moneta, M. V., et al. (2021b). Determinants of subjective wellbeing trajectories in older adults: A growth mixture modeling approach. *Journal of Happiness Studies, 22*(2), 709–726. https://doi.org/10.1007/s10902-020-00248-2

Moreno-Agostino, D., Stone, A. A., Schneider, S., Koskinen, S., Leonardi, M., Naidoo, N., et al. (2020b). Are retired people higher in experiential wellbeing than working older adults? A time use approach. *Emotion, 20*(8), 1411–1422. https://doi.org/10.1037/emo0000637

Mulder, J. D., & Hamaker, E. L. (2020). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling: A Multidisciplinary Journal*. https://doi.org/10.1080/10705511.2020.1784738

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide*, 8th edn.

Ngamaba, K. H., Panagioti, M., & Armitage, C. J. (2017). How strongly related are health status and subjective well-being? Systematic review and meta-analysis. *European Journal of Public Health, 27*(5), 879–885. https://doi.org/10.1093/eurpub/ckx081

OECD. (2013). *OECD guidelines on measuring subjective well-being*. OECD Publishing.

Pavot, W., & Diener, E. (2008). The satisfaction with life scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology, 3*(2), 137–152. https://doi.org/10.1080/17439760701756946

Puvill, T., Kusumastuti, S., Lund, R., Mortensen, E. L., Slaets, J., Lindenberg, J., et al. (2019). Do psychosocial factors modify the negative association between disability and life satisfaction in old age? *PLoS ONE, 14*(10), e0224421. https://doi.org/10.1371/journal.pone.0224421

Puvill, T., Lindenberg, J., de Craen, A. J., Slaets, J. P., & Westendorp, R. G. (2016). Impact of physical and mental health on life satisfaction in old age: A population based observational study. *BMC Geriatrics, 16*(1), 194. https://doi.org/10.1186/s12877-016-0365-4

R Core Team. (2020). *R: A language and environment for statistical computing* (3.6.3). R Foundation for Statistical Computing.

Robitzsch, A. (2020). *sirt: Supplementary item response theory models*.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Salomon, J. A., Mathers, C. D., Chatterji, S., Sadana, R., Ustun, T. B., & Murray, C. J. (2003). Quantifying individual levels of health: Definitions, concepts and measurement issues. In C. Murray & D. Evans

(Eds.), *Health systems performance assessment: Debates, methods and empiricism* (pp. 301–318). World Health Organization.

StataCorp, . (2015). *Stata statistical software: Release 14*. StataCorp LP.

Steptoe, A., Breeze, E., Banks, J., & Nazroo, J. (2013). Cohort profile: The English longitudinal study of ageing. *International Journal of Epidemiology, 42*(6), 1640–1648. https://doi.org/10.1093/ije/dys168

Steptoe, A., Deaton, A., & Stone, A. A. (2015). Subjective wellbeing, health, and ageing. *Lancet, 385*(9968), 640–648. https://doi.org/10.1016/S0140-6736(13)61489-0

Steptoe, A., & Fancourt, D. (2019). Leading a meaningful life at older ages and its relationship with social engagement, prosperity, health, biology, and time use. *Proceedings of the National Academy of Sciences, 116*(4), 1207–1212. https://doi.org/10.1073/pnas.1814723116

United Nations. (2017). *World population ageing report 2017*. United Nations.

Usami, S. (2020). On the differences between general cross-lagged panel model and random-intercept cross-lagged panel model: Interpretation of cross-lagged parameters and model choice. *Structural Equation Modeling: A Multidisciplinary Journal*. https://doi.org/10.1080/10705511.2020.1821690

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods, 24*(5), 637–657. https://doi.org/10.1037/met0000210

Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review, 96*(2), 234–254. https://doi.org/10.1037/0033-295x.96.2.234

World Health Organization. (2015). *World report on ageing and health*. World Health Organization.

World Health Organization. (2017). *Global strategy and action plan on ageing and health*. World Health Organization.

Zyphur, M. J., Voelkle, M. C., Tay, L., Allison, P. D., Preacher, K. J., Zhang, Z., et al. (2019). From data to causes II: Comparing approaches to panel data analysis. *Organizational Research Methods*. https://doi.org/10.1177/1094428119847280

Authors and Affiliations

Darío Moreno-Agostino[1,2,3] · **Francisco José Abad[4]** · **Francisco Félix Caballero[5,6]**

[1] Department of Health Service and Population Research, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

[2] Centre for Longitudinal Studies, UCL Social Research Institute, University College London, London, UK

[3] ESRC Centre for Society and Mental Health, King's College London, London, UK

[4] Department of Social Psychology and Methodology, School of Psychology, Universidad Autónoma de Madrid, Madrid, Spain

[5] Department of Preventive Medicine, Public Health, and Microbiology, School of Medicine, Universidad Autónoma de Madrid, Madrid, Spain

[6] Centre for Biomedical Research in Epidemiology and Public Health (CIBERESP), Instituto de Salud Carlos III, Madrid, Spain