Title

The mediating role of combined lifestyle factors on the relationship between education and gastric cancer in the Stomach cancer Pooling (StoP) Project

Journal

British Journal of Cancer, 127(5)

ISSN

0007-0920

Authors

Alicandro, Gianfranco
Bertuccio, Paola
Collatuzzo, Giulia
et al.

Publication Date

2022-09-01

DOI

10.1038/s41416-022-01857-9

Peer reviewed

British Journal of Cancer www.nature.com/bjc

ARTICLE

Epidemiology

The mediating role of combined lifestyle factors on the relationship between education and gastric cancer in the Stomach cancer Pooling (StoP) Project

Gianfranco Alicandro[1,2,25]✉, Paola Bertuccio[3,4,25], Giulia Collatuzzo[5], Claudio Pelucchi[4], Rossella Bonzi[4], Linda M. Liao[6], Charles S. Rabkin[6], Rashmi Sinha[6], Eva Negri[4,5,7], Michela Dalmartello[4], David Zaridze[8], Dmitry Maximovich[8], Jesus Vioque[9,10], Manoli Garcia de la Hera[9,10], Shoichiro Tsugane[11,12], Akihisa Hidaka[11], Gerson Shigueaki Hamada[13], Lizbeth López-Carrillo[14], Raúl Ulises Hernández-Ramírez[15], Reza Malekzadeh[16], Farhad Pourfarzi[17], Zuo-Feng Zhang[18], Robert C. Kurtz[19], M. Constanza Camargo[6], Maria Paula Curado[20], Nuno Lunet[21,22,23], Paolo Boffetta[5,24] and Carlo La Vecchia[4]

BACKGROUND: The causal pathway between high education and reduced risk of gastric cancer (GC) has not been explained. The study aimed at evaluating the mediating role of lifestyle factors on the relationship between education and GC

METHODS: Ten studies with complete data on education and five lifestyle factors (smoking, alcohol drinking, fruit and vegetable intake, processed meat intake and salt consumption) were selected from a consortium of studies on GC including 4349 GC cases and 8441 controls. We created an a priori score based on the five lifestyle factors, and we carried out a counterfactual-based mediation analysis to decompose the total effect of education on GC into natural direct effect and natural indirect effect mediated by the combined lifestyle factors. Effects were expressed as odds ratios (ORs) with a low level of education as the reference category.

RESULTS: The natural direct and indirect effects of high versus low education were 0.69 (95% CI: 0.62–0.77) and 0.96 (95% CI: 0.95–0.97), respectively, corresponding to a mediated percentage of 10.1% (95% CI: 7.1–15.4%). The mediation effect was limited to men.

CONCLUSIONS: The mediation effect of the combined lifestyle factors on the relationship between education and GC is modest. Other potential pathways explaining that relationship warrants further investigation.

British Journal of Cancer (2022) 127:855–862; https://doi.org/10.1038/s41416-022-01857-9

BACKGROUND

Gastric cancer (GC) incidence and mortality have steadily declined over the last 50 years [1]. Nevertheless, GC remains an important contributor to the global burden of cancer being the fifth most common cancer and the fourth cause of cancer mortality worldwide [2].

Lifestyle factors play an important role in the development of GC. Smoking and heavy alcohol drinking have been associated with about 30% increased risk of GC [3, 4], and excess risks have been also reported for processed meat [5] and high salt intakes [6]. On the other hand, high consumption of fruit and vegetables is associated with reduced risk [7, 8].

[1]Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy. [2]Cystic Fibrosis Center, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy. [3]Department of Public Health, Experimental and Forensic Medicine, University of Pavia, Pavia, Italy. [4]Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milan, Italy. [5]Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy. [6]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA. [7]Pegaso Online University, Naples, Italy. [8]Department of Clinical Epidemiology, N.N. Blokhin National Medical Research Center for Oncology, Moscow, Russia. [9]Instituto de Investigación Sanitaria y Biomédica de Alicante, Universidad Miguel Hernandez (ISABIAL-UMH), Alicante, Spain. [10]Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), Madrid, Spain. [11]Epidemiology and Prevention Group, Center for Public Health Sciences, National Cancer Center, Tokyo, Japan. [12]National Institute of Health and Nutrition, National Institutes of Biomedical Innovation, Health and Nutrition, Tokyo, Japan. [13]Nikkei Disease Prevention Center, São Paulo, Brazil. [14]Mexico National Institute of Public Health, Morelos, Mexico. [15]Department of Biostatistics, Yale School of Public Health, Yale School of Medicine, New Haven, CT, USA. [16]Digestive Oncology Research Center, Digestive Disease Research Institute, Tehran University of Medical Sciences, Tehran, Iran. [17]Digestive Disease Research Center, Ardabil University of Medical Sciences, Ardabil, Iran. [18]Department of Epidemiology, UCLA Fielding School of Public Health and Jonsson Comprehensive Cancer Center, Los Angeles, CA, USA. [19]Department of Medicine, Memorial Sloan Kettering Cancer Centre, New York, NY, USA. [20]Centro Internacional de Pesquisa, A. C. Camargo Cancer Center, São Paulo, Brasil. [21]EPIUnit—Instituto de Saúde Pública da Universidade do Porto, Porto, Portugal. [22]Laboratório para a Investigação Integrativa e Translacional em Saúde Populacional (ITR), Porto, Portugal. [23]Departamento de Ciências da Saúde Pública e Forenses e Educação Médica, Faculdade de Medicina da Universidade do Porto, Porto, Portugal. [24]Stony Brook Cancer Center, Stony Brook University, Stony Brook, NY, USA. [25]These authors contributed equally: Gianfranco Alicandro, Paola Bertuccio. ✉email: gianfranco.alicandro@unimi.it

Received: 4 January 2022 Revised: 2 May 2022 Accepted: 10 May 2022
Published online: 27 May 2022

Most studies assessing the impact of lifestyle on cancer risk have considered each factor individually, but this approach does not capture their complex relationship, where factors act simultaneously showing combined effects [9–11]. To better describe this complex relationship, some authors have used lifestyle scores resulting from the combination of multiple factors [12–14].

GC is among the cancers showing the highest level of socioeconomic disparities, with risk ratios between the lowest and the highest socioeconomic status (SES) around two [15–17]. This social disadvantage has long been known, but the causal pathway between SES and GC has not been fully explained. Lifestyle factors are unevenly distributed among socioeconomic strata with individuals of low SES having, in many populations, less healthy lifestyles [18]. However, the role of lifestyle risk factors as potential mediators of the relationship between low SES and GC remains largely unexplored.

Education is frequently used as a proxy of SES since it captures, besides income, the knowledge assets of the individual and thus the ability to make healthy choices. It is often preferred to other SES indicators because it is relatively easy to collect, it does not usually change over the life course and can be used at all ages and in both sexes irrespective of the employment status.

The aim of this study is to quantify the mediated effect of selected lifestyle factors, including tobacco smoking, heavy alcohol drinking and intake of fruit, vegetables, processed meat and salt on the relationship between education and GC.

METHODS

Participants

We used data from an international consortium of 34 observational studies on GC, the "Stomach Cancer Pooling (StoP) Project", including individual data from 13,121 GC cases and 31,420 controls (version 3.2 of the StoP database) [19]. Potentially relevant epidemiological studies were identified through a literature search, and principal investigators were invited to participate and share original patient-level data. All collected data were harmonised according to a standard format at the coordinating centre and subject identifiers were removed before data pooling. Detailed information about the collection and harmonisation of data in the StoP consortium is given elsewhere [19].

Analyses were based on 10 studies (nine case-control studies and one case-control study nested within a cohort) with complete data on education and five selected lifestyle factors. This subset of studies included two studies from Italy [20, 21], one from Russia [22], one from Spain [23], one from Iran [24], two from the USA [25, 26], one from Mexico [27] and two from Brazil [28, 29]. Six studies used hospital controls [20–23, 25, 28], three studies [24, 26, 27] used population controls and one study used a combination of hospital and population controls [29]. Twenty-four studies were excluded: three studies did not collect data on education, 20 did not collect data on all the lifestyle factors needed for the analysis and one was excluded since it had more than 10% of missing values for one of the lifestyle factors. The study population consisted of 4349 cases and 8441 controls.

Exposure

The exposure variable was the highest attained level of education as reported in the original studies. We adopted a study-specific classification of education in three levels, i.e. low, intermediate and high. This is because the analysis was based on data from studies conducted in different periods, between mid-1980's and the mid-2000's, and in different social and economic contexts. Thus using a study-specific classification allowed to take into account temporal improvement in educational attainment and differences in educational opportunities across countries [30].

Mediators

The mediators of our analysis were five lifestyle factors, including cigarette smoking (categorised as "never smokers", "former smokers", "≤10 cigarettes per day", "11–20 cigarettes per day" and ">20 cigarettes per day"), alcohol intake ("never or ≤12 g of ethanol per day", "13–47 g per day" and ">47 g per day"), consumption of fruit and vegetables ("low", "intermediate" and "high" according to study-specific tertiles), consumption of processed meat ("low", "intermediate" and "high" according to study-specific tertiles) and consumption of salt ("low", "intermediate" and "high" according to study-specific tertiles).

To obtain an overall measure of the combined mediated effect of lifestyle factors on the relationship between education and GC risk, we created an a priori healthy lifestyle score by assigning points to study participants according to their risk factor profile (Supplementary Table 1). The final score was then computed by adding up all the points obtained for each item of the score. This algorithm gave a score ranging from 0 to 12, with higher scores indicating a "healthier lifestyle". Subjects were then grouped into categories based on sex-specific tertiles of the score (i.e. 0–5, 6–8, 9–12 points among men and 0–7, 8–10, 11–12 points among women).

Statistical analysis

We estimated the pooled odds ratios (ORs) for GC and their 95% confidence intervals (CIs) according to education levels and lifestyle factors (considered individually or combined in a score) through mutually adjusted logistic regression models, using a one-stage approach [31]. The models included sex, age (continuous) and study as covariates. ORs for single studies were also computed and given in the Supplementary Information.

To quantify the mediating role of combined lifestyle factors on the relationship between education and GC risk, we carried out a counterfactual mediation analysis using an imputation-based method, which requires the expansion of the data and the imputation of the unobserved counterfactuals by fitting a model for the outcome conditional on the exposure, the mediator and the covariates [32]. The directed acyclic graph (DAG) which depicts the hypothesised causal model is shown in Fig. 1. This method allows to decompose the total effect (TE) of education on GC risk into a natural indirect effect (NIE) through the lifestyle score and the remaining natural direct effect (NDE). When the NIE was statistically significant, the percentages mediated were estimated as (NIE/TE) · 100. The NDE indicates the average effect of the exposure (i.e. intermediate or high education) on GC risk, when the mediator is kept at the level it would have taken in the absence of exposure (i.e. low education). The NIE is defined as the average effect of the exposure when the mediator is set to the level it would have been with versus without exposure. NDE and NIE were each estimated by logistic regression models adjusted for study, sex and age. They were expressed as ORs with their 95% CIs, with the latter computed using robust standard errors based on a sandwich estimator. The 95% CIs for the percentages mediated were obtained by bootstrapping 1000 samples.

We further investigated through a moderated mediation analysis whether the mediated effects differed across strata of sex, age group (<65 and ≥65), and geographic area (Europe, North America, Latin America, Asia) [33]. The likelihood ratio test was used for testing the statistical significance of the moderated mediation effects.

We handled missing data by applying a multivariate imputation by chained equations [34]. We imputed missing data five times, thus generating five imputed datasets that were used for the analysis. To impute missing values for education and lifestyle scores we used sex, age

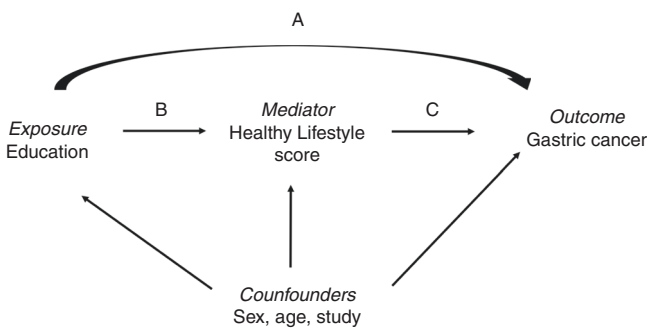

Fig. 1 Directed acyclic graph showing the relationship between education and gastric cancer risk and the decomposition of the effects. Arrow A displays the natural direct effect (NDE) of education on gastric cancer risk, while path B + C displays the natural indirect effect (NIE) mediated by the healthy lifestyle score. The sum of NDE and NIE gives the total effect (TE). The last three arrows display the confounding variables.

and study as predictors. The models were then fitted to each imputed dataset separately, and the parameter and variance estimates obtained were combined using the Rubin's rule [35], i.e. the coefficients of the regression models estimated from each imputed dataset were averaged to obtain a pooled estimate with CIs including the uncertainty introduced by the imputation of missing values.

Four sensitivity analyses were carried out: (1) we added two interaction terms to the mediation models, i.e. sex-by-study and age-by-study, to account for possible imbalance in the distribution of sex and age within the included studies; (2) we excluded the studies where education was not associated with GC; (3) we analysed separately the nested case-control study and the remaining studies; (4) we analysed separately studies using hospital and population controls.

To carry out the mediation analysis we used the "neImpute" function from the R package "medflex" [33], while the "mice" function from the "mice" package was used for the multiple imputation procedure [34].

RESULTS

Study subjects

Table 1 shows the distribution of 4349 GC cases and 8441 controls according to study, sex, age, education and the five lifestyle factors considered individually or combined. Most subjects were from Europe (43%) and the USA (40%), 11% were from Latin America and around 5% from Iran. About 32% of cases and 35% of controls were females. Cases were older, less educated and more frequently heavy smokers and heavy alcohol drinkers than controls. As for dietary factors, cases reported similar consumption of fruit and vegetables and higher consumption of processed meat and salt as compared to controls. The distribution of the healthy lifestyle score showed a higher proportion of cases than controls in the lowest score category. Missing values for study variables were below 3% in both cases and controls.

Lifestyle factors across levels of education

Figure 2 shows the distribution of the lifestyle factors across levels of education. The percentages of heavy smokers were 8.5%, 5.8% and 4.1% among low, intermediate and high education levels, while the corresponding figures for heavy alcohol drinkers were 20.5%, 12.3% and 13.1%. Only negligible differences were observed for the other lifestyle factors. When single factors were combined in the lifestyle score, the percentages of individuals with values below the 1st tertile were 41.4%, 35.3% and 31.4% among low, intermediate and high education levels, respectively.

Associations between education, lifestyle factors and GC

The pooled associations between education, lifestyle factors (considered individually or combined) and GC are given in Table 2. As compared to subjects with a low education level, the ORs were 0.78 (95% CI: 0.70–0.86) for those with intermediate education and 0.70 (95% CI: 0.62–0.78) for those with high education. For smoking habits, taking as a reference more than 20 cigarettes per day, the ORs were below unity for all the other categories, with an OR of 0.66 (95% CI: 0.56–0.78) for never smokers. Compared to heavy drinkers (>47 g of ethanol/day), the ORs were 0.90 for moderate drinkers (13–47 g/day) and 0.88 for never/low drinkers (≤12 g/day), although the CIs included unity. Inverse associations with GC emerged for high vs low fruit and vegetable intake (OR 0.77, 95% CI 0.70–0.85) as well as for low vs high consumption of processed meat (OR 0.77, 95% CI 0.70–0.84) and low salt intake (OR 0.80, 95% CI 0.72–0.89).

An inverse association was also found between GC and the lifestyle score. As compared to the first category, which indicates a less healthy lifestyle, the ORs were 0.76 (95% CI 0.69–0.83) for the intermediate and 0.65 (95% CI 0.58–0.72) for the highest category.

The associations obtained in each study separately are shown in Supplementary Table 2. The point estimates of the ORs indicate an inverse association between education and GC in all studies except those from Spain [23] and Mexico [27], although some CIs

Table 1. Distribution of the 4349 gastric cancer cases and 8441 controls according to selected covariates within the Stomach Cancer Pooling (StoP) Project.

	Controls		Cases	
	N	%	*N*	%
Study				
Europe				
Italy 1 [20]	2081	23.3	769	17.0
Italy 2 [21]	547	6.1	230	5.1
Russia [22]	611	6.8	450	10.0
Spain 2 [23]	455	5.1	401	8.9
Asia				
Iran 1 [24]	394	4.4	217	4.8
North America				
USA 1 [25]	132	1.5	132	2.9
USA 4 [26]	3331	37.2	1583	35.0
Latin America				
Mexico 1 [27]	478	5.3	248	5.5
Brazil 1 [28]	226	2.5	226	5.0
Brazil 2 [29]	186	2.1	93	2.1
Sex				
Women	2932	34.7	1405	32.3
Men	5509	65.3	2944	67.7
Age				
<40	447	5.3	109	2.5
40–44	370	4.4	113	2.6
45–49	504	6.0	215	4.9
50–54	618	7.3	288	6.6
55–59	897	10.6	458	10.5
60–64	1129	13.4	632	14.5
65–69	1609	19.1	930	21.4
70–74	1530	18.1	885	20.3
≥75	1337	15.8	719	16.5
Level of education				
Low	2703	32.0	1766	40.6
Intermediate	3282	38.9	1554	35.7
High	2355	27.9	949	21.8
Missing	101	1.2	80	1.8
Cigarette smoking				
Never smokers	3682	43.6	1698	39.0
Former smokers	2803	33.2	1565	36.0
0 < cigarettes per day ≤ 10	559	6.6	239	5.5
10 < cigarettes per day ≤ 20	773	9.2	409	9.4
>20 cigarettes per day	455	5.4	329	7.6
Missing	169	2.0	109	2.5
Alcohol intake (grams of ethanol per day)				
Never drinkers	2954	35.0	1527	35.1
Low (≤12 g)	2502	29.6	1199	27.6
Intermediate (>12 and ≤47 g)	1747	20.7	870	20.0
High (>47 g)	1217	14.4	729	16.8
Missing	21	0.2	24	0.6
Fruit and vegetable intake				
Low	2255	26.7	1225	28.2
Intermediate	2835	33.6	1463	33.6
High	3343	39.6	1652	38.0
Missing	8	0.1	9	0.2

Table 1. continued

	Controls		Cases	
	N	%	*N*	%
Processed meat intake				
Low	3272	38.8	1497	34.4
Intermediate	2562	30.4	1327	30.5
High	2597	30.8	1513	34.8
Missing	10	0.1	12	0.3
Salt intake				
Low	3422	40.5	1642	37.8
Intermediate	3009	35.6	1621	37.3
High	2000	23.7	1080	24.8
Missing	10	0.1	6	0.1
Healthy lifestyle score category (based on tertiles)				
1	2853	33.8	1677	38.6
2	3031	35.9	1446	33.2
3	2355	27.9	1087	25.0
Missing	202	2.4	139	3.2

included unity likely due to the limited numbers of study subjects. The lifestyle score was inversely associated with GC in all studies except the one from Mexico [27] and a study from Italy [21]. Again, some CIs included unity. Due to the low number of individuals in several categories, results for individual lifestyle factors were unstable and were not reported.

Mediation of lifestyle factors on the relationship between education and GC risk

The a priori lifestyle score explained about 5.6% (95% CI: 2.6–10.6) of the association between intermediate vs low education and GC, and 10.1% (95% CI: 7.1–15.4) of that between high *vs* low education and GC, as indicated by the results of the mediation analysis summarised in Table 3.

Figure 3 shows the results of the moderated mediation analysis in strata of sex, age and geographic area. The test for the moderated effect indicated that the NIEs were significantly different among strata of sex ($p < 0.001$, Fig. 3e, f), whereas no significant differences were observed among strata of age (<65 vs ≥65 years) and geographic area. Lifestyle score partly mediated the differences in GC risk between education levels only among men, with percentages mediated of 8.0% (95% CI: 3.9–15.0) for

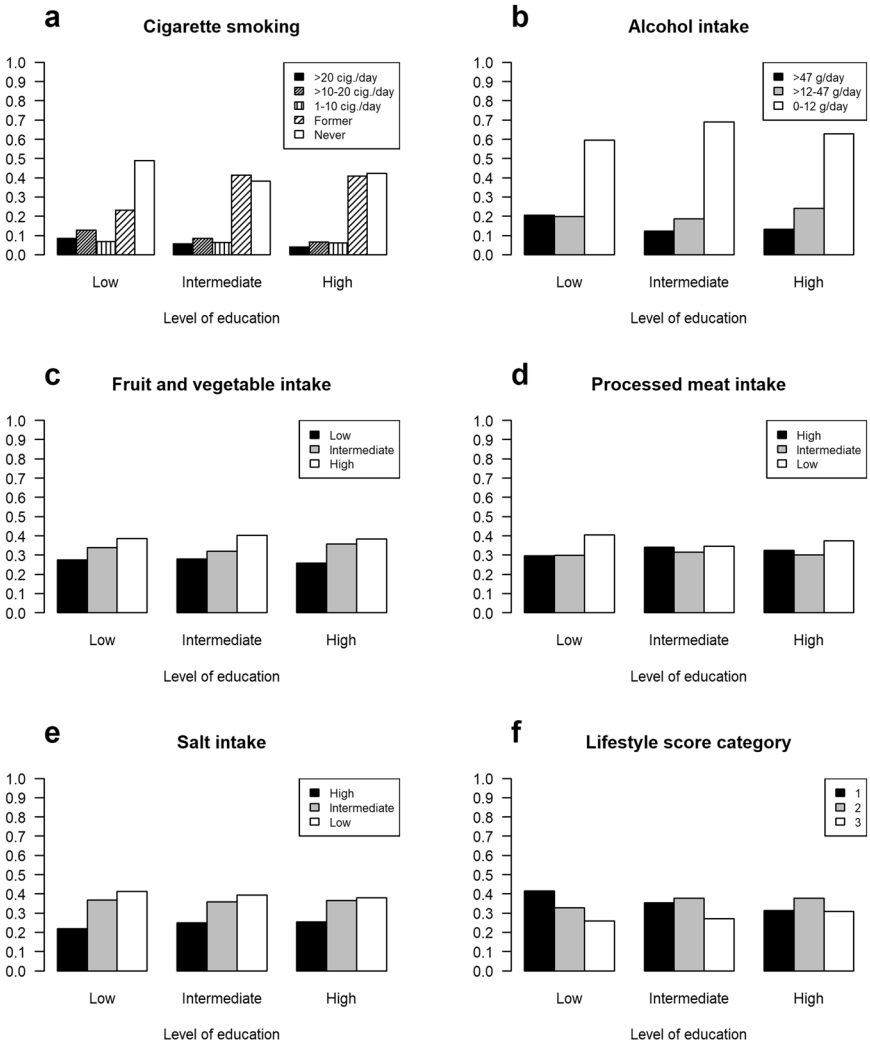

Fig. 2 Lifestyle factors across levels of education. Distribution of cigarette smoking (**a**), alcohol intake (**b**), fruit and vegetable intake (**c**), processed meat intake (**d**), salt intake (**e**) and healthy lifestyle score categories (**f**). Intakes of fruit and vegetable, processed meat and salt were categorized according to study-specific levels. Categories of lifestyle score were obtained from sex-specific tertiles.

intermediate and 10.7% (95% CI: 6.8–16.4) for high vs low education (Supplementary Table 3).

Results of the sensitivity analyses are reported in Supplementary Tables 4–7. The inclusion of the interaction terms (sex-by-study and age-by-study), as well as the exclusion of the studies where education was not associated with GC, did not materially change the results of the main analysis. A lower mediation effect was found in studies based on population controls as compared to hospital controls (percentage mediated: 5.1 vs 12.6 for the high level of education) as well as in the nested case-control study as compared to the remaining studies (percentage mediated: 6.2 vs 12.8% for the high level of education).

Table 2. Odds ratios for gastric cancer and 95% confidence intervals according to levels of education and selected lifestyle factors.

	OR	95% CI
Model 1		
Level of education (ref: Low)		
Intermediate	0.78	(0.70–0.86)
High	0.70	(0.62–0.78)
Cigarette smoking (ref: >20 cigarettes per day)		
10 < cigarettes per day ≤ 20	0.78	(0.65–0.95)
0 < cigarettes per day ≤ 10	0.68	(0.55–0.84)
Former smokers	0.81	(0.68–0.95)
Never smokers	0.66	(0.56–0.78)
Alcohol intake (grams of ethanol per day) (ref: High, i.e. >47 g)		
Intermediate (>12 - ≤47 g)	0.90	(0.79–1.03)
Never and low drinkers (≤12 g)	0.88	(0.78–1.00)
Fruit and vegetable intake (ref: Low)		
Intermediate	0.91	(0.82–1.00)
High	0.77	(0.70–0.85)
Processed meat intake (ref: High)		
Intermediate	0.85	(0.77–0.94)
Low	0.77	(0.70–0.84)
Salt intake (ref: High)		
Intermediate	1.05	(0.95–1.17)
Low	0.80	(0.72–0.89)
Model 2		
Level of education (ref: Low)		
Intermediate	0.78	(0.70–0.86)
High	0.69	(0.61–0.77)
Healthy lifestyle score category (based on tertiles) (ref: 1)		
2	0.76	(0.69–0.83)
3	0.65	(0.58–0.72)

OR odds ratio, *CI* confidence intervals.
ORs were estimated using two different logistic regression models: Model 1 included education, all lifestyle factors, sex, age and study as covariates; Model 2 included education, the lifestyle score (obtained from the combination of the individual factors) and the same set of covariates of Model 1.

DISCUSSION

The combination of the main lifestyle risk factors for GC, namely tobacco, alcohol, low consumption of fruit and vegetables, high consumption of processed meat and salt intake, mediate only about 10% of the difference in GC risk between highly and less educated individuals. The mediation effect was restricted to men.

Our findings indicate therefore that different pathways are in place, which, beyond lifestyle behaviours, link low education to increased risk of GC. One of such pathways may involve chronic atrophic gastritis caused by *Helicobacter pylori* (*Hp*) infection, which is strongly associated with GC [36] and more frequently reported among people of low SES [37, 38] and in countries with less developed economic and health systems [39–41]. Several socioeconomic factors are associated with the transmission of the infection at different ages, such as low family income, low education, living in rural areas, in crowded places with no tap water and toilet outside the house [39]. Thus, SES predisposes to *Hp* infection in many ways, mostly acting in early childhood. *Hp* infection may also interact with lifestyle factors further widening the socioeconomic disparities in GC [37].

Another possible explanation of the limited mediation effect found in our study is the possible role of residence and birthplace. In fact, people living and children born in rural areas are commonly at higher risk of developing GC compared to people from urban areas [42], possibly due to higher exposures to *Hp* occurring early in life, whose effect extends into adulthood.

In addition, less educated individuals could be more exposed to environmental carcinogens (such as dust, fumes and selected chemicals), both occupationally and residentially than highly educated individuals [43, 44].

Time at diagnosis can be also a potential mediator of the SES disparity in GC incidence. In particular, education can be associated with the detection and control of preneoplastic lesions, because of more attention to symptoms and generally better personal healthcare.

We cannot directly compare our results with other published data since, to our knowledge, no previous studies quantified the contribution of lifestyle factors to educational inequalities in GC aetiology. Some attempts have been made on other cancer sites. A study on colorectal cancer found that lifestyle behaviours (i.e.

Table 3. Mediation effects of the lifestyle score on the relationship between education and gastric cancer.

	OR	95% CI	*p*-value	% mediated (95% CI)
Education (ref: Low)				
Intermediate				
Natural direct effect	0.778	(0.703–0.860)	<0.001	
Natural indirect effect	0.985	(0.977–0.994)	0.001	5.6 (2.6–10.6)
Total effect	0.766	(0.692–0.848)	<0.001	
High				
Natural direct effect	0.690	(0.617–0.772)	<0.001	
Natural indirect effect	0.959	(0.947–0.972)	<0.001	10.1 (7.1–15.4)
Total effect	0.662	(0.592–0.740)	<0.001	

OR odds ratio, *CI* confidence intervals.

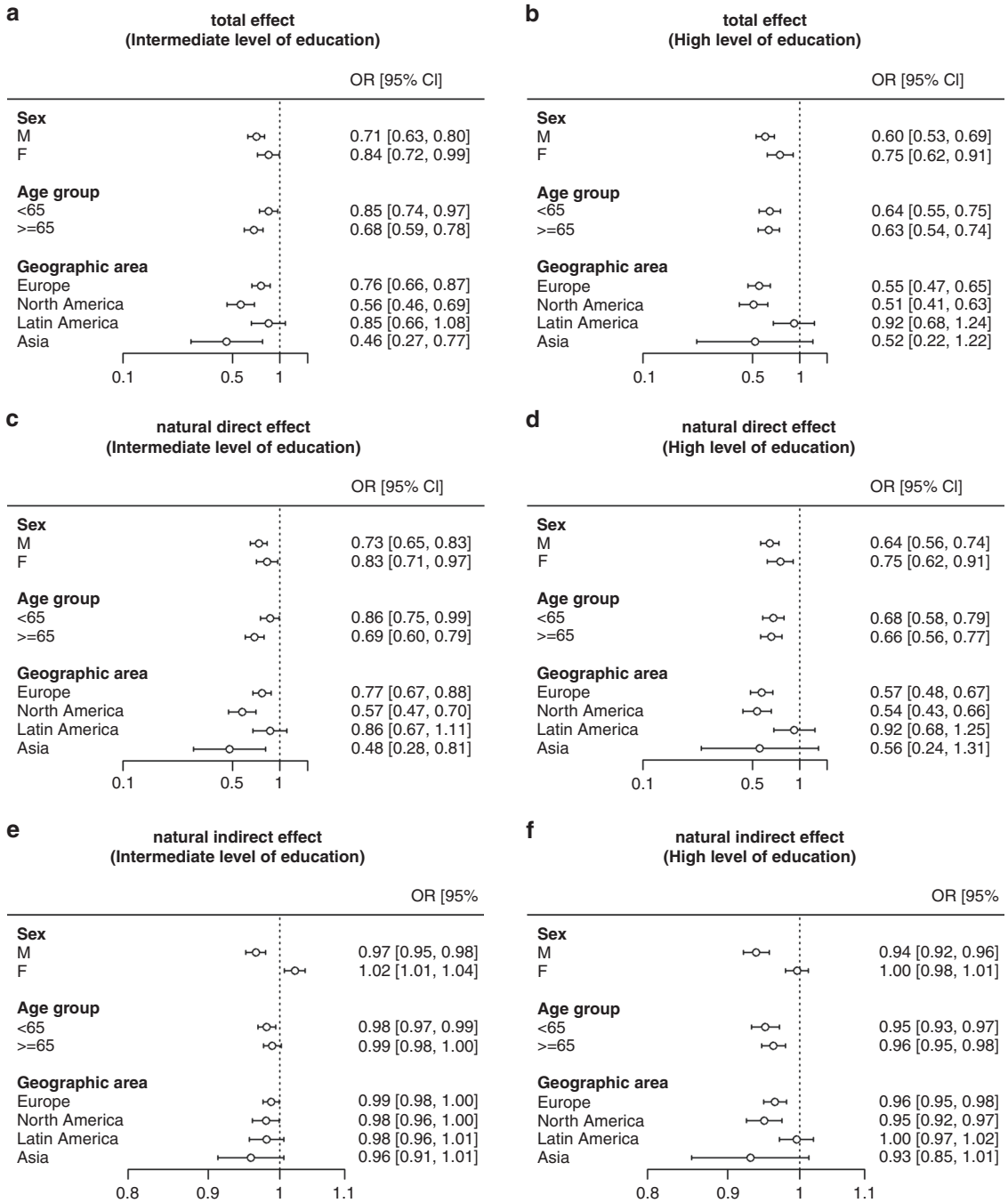

Fig. 3 Results of the moderated mediation analysis: mediation effects of lifestyle score on the relationship between education and gastric cancer moderated by sex, age group and geographic area. Total effect for intermediate (**a**) and high level (**b**) of education. Natural direct effect (NDE) for intermediate (**c**) and high level (**d**) of education. Natural indirect effect (NIE) for intermediate (**e**) and high level (**f**) of education. Test of interactions for moderated mediation effects: (1) for sex, $p = 0.274$ for NDE and $p < 0.001$ for NIE; (2) for age group: $p = 0.034$ for NDE and $p = 0.623$ for NIE; (3) for geographic area, $p = 0.012$ for NDE and $p = 0.140$ for NIE.

smoking history, Mediterranean diet and physical activity) explained almost 40% of the difference in colorectal cancer risk between high and low educational attainment, with the Mediterranean diet mediating 22% of the difference [45].

Considering the five lifestyle factors together in a composite score in relation to GC, we found that healthy behaviours were independently associated with a 35% risk reduction of GC. This estimate is in line with the findings of a study based on the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, which found a 51% lower risk of GC in the highest category of a healthy lifestyle index based on smoking status, alcohol consumption and a diet quality score reflecting the Mediterranean diet [12]. Similar findings were observed in a Chinese study with a hazard ratio for GC of 0.42 (95% CI: 0.31–0.57) for the highest compared to the lowest category of a healthy score based on tobacco, alcohol, obesity and dietary habits [46].

In our study, results from the moderated mediation analysis suggest that the relationship between education level and GC risk

was partly mediated by lifestyle factors only among men. The lack of mediation effect among women could be largely attributable to the healthier lifestyle of women than men [47, 48]. Moreover, highly educated women tended to have less healthy behaviour than their low educated counterparts (Supplementary Table 8).

When interpreting our results, it should be noted that the patterns of smoking, alcohol drinking and other dietary factors across educational levels may have changed by sex and country, and also over time. Similarly, educational attainment may have improved at a different pace depending on the socioeconomic context of the country. Moreover, the magnitude of the association between education and GC may vary according to the country where data were collected. For example, in a previous study of this consortium which included a larger set of studies from different countries [17], the OR between low education and GC was 0.60 as compared to 0.70 we found in the present study. This difference may also apply to the estimate of the mediating role of the five lifestyle factors.

Among potential study limitations, we considered education as a proxy of SES for all the studies included in the analysis, independently from the level of socioeconomic development of the country. In studies from less wealthy countries, other indicators, not available for the majority of the included studies, may better reflect the socioeconomic stratification of the population, such as disposable income, household conditions (presence of damp, building materials, number of rooms and overcrowding) and facilities potentially linked to GC (access to hot and cold water, heating, sole use of bathrooms and toilets, whether the toilet is inside or outside the house, having a refrigerator, washing machine or telephone) [30, 49]. However, the validity of education as SES indicator has been widely described, suggesting that a healthy lifestyle is promoted by knowledge and self-awareness besides economic conditions [43, 48]. Second, information about smoking, alcohol and dietary habits may suffer from possible recall bias since they have been collected retrospectively in all the included studies but one [25]. In this latter study, the estimate of the mediation effect was lower than that obtained from the remaining studies. Third, GC cases may have changed their lifestyle habits after the diagnosis. Fourth, most studies enrolled hospital-based controls which could have biased the reported prevalence of lifestyle factors. We tried to estimate the effect of this potential bias by analysing separately the studies having hospital and population controls and we found that the mediation effect of lifestyle factors was smaller in studies enrolling population controls. However, this cannot be entirely attributed to the type of controls since it is based on a selection of different studies and thus other unmeasured factors may account for the observed difference. Finally, the models were not adjusted for *Hp* infection, as the information was available only for four of the included studies (the study from Iran, that from Mexico and the two studies from Brazil) with most subjects (70–80%) infected and no association found between *Hp* infection and GC as well as between education and *Hp* infection (Supplementary Tables 9–10) [32, 50].

The overall impact of all these potential biases is difficult to assess since they could lead to either underestimation or overestimation of the real mediating effect. Despite these potential limitations, our study is a first attempt to quantify the mediating effect of combined lifestyle behaviours on the relationship between education and GC risk. Moreover, we used individual-level data from a global consortium, which gave us the opportunity to analyse a unique dataset with large numbers of cases and controls.

In conclusion, our data show that the combination of the main lifestyle risk factors for GC mediated 10% of the education inequality in GC risk among men, while no mediating effect was detected among women. These findings suggest that other pathways linking education to GC should be considered.

DATA AVAILABILITY

The data that support the findings of our study are available from the Stomach cancer Pooling (StoP) Project but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however, available from the authors upon reasonable request and with permission of the Steering Committee of the StoP Project.

REFERENCES

1. Santucci C, Carioli G, Bertuccio P, Malvezzi M, Pastorino U, Boffetta P, et al. Progress in cancer mortality, incidence, and survival: a global overview. Eur J Cancer Prev. 2020;29:367–81.
2. Ferlay J, Ervik M, Lam F, Comobet M, Mery L, Piñeros M, et al. Global Cancer Observatory: Cancer Today. International Agency for Research on Cancer, 2020. https://gco.iarc.fr/today/home. Accessed 15 April 2021.
3. Praud D, Rota M, Pelucchi C, Bertuccio P, Rosso T, Galeone C, et al. Cigarette smoking and gastric cancer in the Stomach Cancer Pooling (StoP) Project. Eur J Cancer Prev. 2018;27:124–33.
4. Rota M, Pelucchi C, Bertuccio P, Matsuo K, Zhang ZF, Ito H, et al. Alcohol consumption and gastric cancer risk-A pooled analysis within the StoP project consortium. Int J Cancer. 2017;141:1950–62.
5. Ferro A, Rosato V, Rota M, Costa AR, Morais S, Pelucchi C, et al. Meat intake and risk of gastric cancer in the Stomach cancer Pooling (StoP) project. Int J Cancer. 2020;147:45–55.
6. D'Elia L, Rossi G, Ippolito R, Cappuccio FP, Strazzullo P. Habitual salt intake and risk of gastric cancer: a meta-analysis of prospective studies. Clin Nutr. 2012;31:489–98.
7. Bertuccio P, Alicandro G, Rota M, Pelucchi C, Bonzi R, Galeone C, et al. Citrus fruit intake and gastric cancer: The stomach cancer pooling (StoP) project consortium. Int J Cancer. 2019;144:2936–44.
8. Ferro A, Costa AR, Morais S, Bertuccio P, Rota M, Pelucchi C, et al. Fruits and vegetables intake and gastric cancer risk: a pooled analysis within the Stomach cancer Pooling Project. Int J Cancer. 2020;147:3090–101.
9. Hausdorf K, Eakin E, Whiteman D, Rogers C, Aitken J, Newman B. Prevalence and correlates of multiple cancer risk behaviors in an Australian population-based survey: results from the Queensland Cancer Risk Study. Cancer Causes Control. 2008;19:1339–47.
10. Kvaavik E, Batty GD, Ursin G, Huxley R, Gale CR. Influence of individual and combined health behaviors on total and cause-specific mortality in men and women: the United Kingdom health and lifestyle survey. Arch Intern Med. 2010;170:711–8.
11. Schuit AJ, van Loon AJ, Tijhuis M, Ocke M. Clustering of lifestyle risk factors in a general adult population. Prev Med. 2002;35:219–24.
12. Buckland G, Travier N, Huerta JM, Bueno-de-Mesquita HB, Siersema PD, Skeie G, et al. Healthy lifestyle index and risk of gastric adenocarcinoma in the EPIC cohort study. Int J Cancer. 2015;137:598–606.
13. Jiao L, Mitrou PN, Reedy J, Graubard BI, Hollenbeck AR, Schatzkin A, et al. A combined healthy lifestyle score and risk of pancreatic cancer in a large cohort study. Arch Intern Med. 2009;169:764–70.
14. Lohse T, Faeh D, Bopp M, Rohrmann S, Swiss National Cohort Study G. Adherence to the cancer prevention recommendations of the World Cancer Research Fund/American Institute for Cancer Research and mortality: a census-linked cohort. Am J Clin Nutr. 2016;104:678–85.
15. Alicandro G, Frova L, Sebastiani G, El Sayed I, Boffetta P, La Vecchia C. Educational inequality in cancer mortality: a record linkage study of over 35 million Italians. Cancer Causes Control. 2017;28:997–1006.
16. Lagergren J, Andersson G, Talback M, Drefahl S, Bihagen E, Harkonen J, et al. Marital status, education, and income in relation to the risk of esophageal and gastric cancer by histological type and site. Cancer. 2016;122:207–12.
17. Rota M, Alicandro G, Pelucchi C, Bonzi R, Bertuccio P, Hu J, et al. Education and gastric cancer risk—an individual participant data meta-analysis in the StoP project consortium. Int J Cancer. 2020;146:671–81.
18. Zhang YB, Chen C, Pan XF, Guo J, Li Y, Franco OH, et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. BMJ. 2021;373:n604.
19. Pelucchi C, Lunet N, Boccia S, Zhang ZF, Praud D, Boffetta P, et al. The stomach cancer pooling (StoP) project: study design and presentation. Eur J Cancer Prev. 2015;24:16–23.
20. La Vecchia C, D'Avanzo B, Negri E, Decarli A, Benichou J. Attributable risks for stomach cancer in northern Italy. Int J Cancer. 1995;60:748–52.
21. Pelucchi C, Tramacere I, Bertuccio P, Tavani A, Negri E, La Vecchia C. Dietary intake of selected micronutrients and gastric cancer risk: an Italian case-control study. Ann Oncol. 2009;20:160–5.
22. Zaridze D, Borisova E, Maximovitch D, Chkhikvadze V. Alcohol consumption, smoking and risk of gastric cancer: case-control study from Moscow, Russia. Cancer Causes Control. 2000;11:363–71.

23. Santibanez M, Alguacil J, de la Hera MG, Navarrete-Munoz EM, Llorca J, Aragones N, et al. Occupational exposures and risk of stomach cancer by histological type. Occup Environ Med. 2012;69:268–75.
24. Pourfarzi F, Whelan A, Kaldor J, Malekzadeh R. The role of diet and other environmental factors in the causation of gastric cancer in Iran-a population based study. Int J Cancer. 2009;125:1953–60.
25. Zhang ZF, Kurtz RC, Klimstra DS, Yu GP, Sun M, Harlap S, et al. Helicobacter pylori infection on the risk of stomach cancer and chronic atrophic gastritis. Cancer Detect Prev. 1999;23:357–67.
26. Schatzkin A, Subar AF, Thompson FE, Harlan LC, Tangrea J, Hollenbeck AR, et al. Design and serendipity in establishing a large cohort with wide dietary intake distributions: the National Institutes of Health-American Association of Retired Persons Diet and Health Study. Am J Epidemiol. 2001;154:1119–25.
27. Hernandez-Ramirez RU, Galvan-Portillo MV, Ward MH, Agudo A, Gonzalez CA, Onate-Ocana LF, et al. Dietary intake of polyphenols, nitrate and nitrite and gastric cancer risk in Mexico City. Int J Cancer. 2009;125:1424–30.
28. Nishimoto IN, Hamada GS, Kowalski LP, Rodrigues JG, Iriya K, Sasazuki S, et al. Risk factors for stomach cancer in Brazil (I): a case-control study among non-Japanese Brazilians in Sao Paulo. Jpn J Clin Oncol. 2002;32:277–83.
29. Hamada GS, Kowalski LP, Nishimoto IN, Rodrigues JJ, Iriya K, Sasazuki S, et al. Risk factors for stomach cancer in Brazil (II): a case-control study among Japanese Brazilians in Sao Paulo. Jpn J Clin Oncol. 2002;32:284–90.
30. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 1). J Epidemiol Community Health. 2006;60:7–12.
31. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. Stat Med. 2017;36:855–75.
32. Vansteelandt S, Bekaert M, Lange T. Imputation strategies for the estimation of natural direct and indirect effects. Epidemiol Methods. 2012;1. https://doi.org/10.1515/2161-962X.1014.
33. Steen J, Loeys T, Moerkerke B, Vansteelandt S. medflex: an R package for flexible mediation analysis using natural effect models. J Stat Softw. 2017;76:1–46.
34. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw. 2017;45:1–67.
35. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons Inc; 1987.
36. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, et al. Helicobacter pylori infection and the development of gastric cancer. N Engl J Med. 2001;345:784–9.
37. Collatuzzo G, Pelucchi C, Negri E, Lopez-Carrillo L, Tsugane S, Hidaka A, et al. Exploring the interactions between Helicobacter pylori (Hp) infection and other risk factors of gastric cancer: a pooled analysis in the Stomach cancer Pooling (StoP) Project. Int J Cancer. 2021;149:1228–38.
38. Boffetta P. Infection with Helicobacter pylori and parasites, social class and cancer. IARC Sci Publ. 1997;(138):325–9.
39. Laszewicz W, Iwanczak F, Iwanczak B. Seroprevalence of Helicobacter pylori infection in Polish children and adults depending on socioeconomic status and living conditions. Adv Med Sci. 2014;59:147–50.
40. Eusebi LH, Zagari RM, Bazzoli F. Epidemiology of Helicobacter pylori infection. Helicobacter. 2014;19(Suppl 1):1–5.
41. Niv Y, Hazazi R. Helicobacter pylori recurrence in developed and developing countries: meta-analysis of 13C-urea breath test follow-up after eradication. Helicobacter. 2008;13:56–61.
42. Ocana-Riola R, Sanchez-Cantalejo C, Rosell J, Sanchez-Cantalejo E, Daponte A. Socio-economic level, farming activities and risk of cancer in small areas of Southern Spain. Eur J Epidemiol. 2004;19:643–50.
43. Kogevinas M, Pearce N, Susser M, Boffetta P. Social Inequalities and Cancer. IARC Science Publications No 138 Lyon: International Agency for Research on Cancer; 1997.
44. Shah SC, Boffetta P, Johnson KC, Hu J, Palli D, Ferraroni M, et al. Occupational exposures and odds of gastric cancer: a StoP project consortium pooled analysis. Int J Epidemiol. 2020;49:422–34.
45. Doubeni CA, Major JM, Laiyemo AO, Schootman M, Zauber AG, Hollenbeck AR, et al. Contribution of behavioral risk factors and obesity to socioeconomic differences in colorectal cancer incidence. J Natl Cancer Inst. 2012;104:1353–62.
46. Wang Z, Koh WP, Jin A, Wang R, Yuan JM. Composite protective lifestyle factors and risk of developing gastric adenocarcinoma: the Singapore Chinese Health Study. Br J Cancer. 2017;116:679–87.
47. Vari R, Scazzocchio B, D'Amore A, Giovannini C, Gessani S, Masella R. Gender-related differences in lifestyle may affect health status. Ann Ist Super Sanita. 2016;52:158–66.
48. Garcia-Mayor J, Moreno-Llamas A, la Cruz-Sanchez E. High educational attainment redresses the effect of occupational social class on health-related lifestyle: findings from four Spanish national health surveys. Ann Epidemiol. 2021;58:29–37.
49. La Vecchia C, Negri E, D'Avanzo B, Franceschi S. Electric refrigerator use and gastric cancer risk. Br J Cancer. 1990;62:136–7.
50. Ferro A, Morais S, Pelucchi C, Aragones N, Kogevinas M, Lopez-Carrillo L, et al. Smoking and Helicobacter pylori infection: an individual participant pooled analysis (Stomach Cancer Pooling- StoP Project). Eur J Cancer Prev. 2019;28:390–6.

ACKNOWLEDGEMENTS

We thank the European Cancer Prevention (ECP) Organization for providing support for the project meetings.

AUTHOR CONTRIBUTIONS

Conceptualisation: GA, PB, GC, PB; Methodology: GA, PB, GC, PB; Data collection: LML, CSR, EN, DZ, JV, ST, GSH, LL-C, RUH-R, RM, Z-FZ, CLV; Formal analysis and investigation: GA, PB; Writing—original draft preparation: GA, PB; Writing—review and editing: All authors; Funding acquisition: CLV; Resources: CLV; Supervision: CLV, CP.

FUNDING

This study was funded by the Associazione Italiana per la Ricerca sul Cancro (AIRC), Project no. 21378 (Investigator Grant).

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1964 and later versions. The StoP Project received ethical approval from the University of Milan Institutional Review Board (reference no. 19/15 of 01/04/2015). Informed consent was obtained for each subject included in the study.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41416-022-01857-9.

Correspondence and requests for materials should be addressed to Gianfranco Alicandro.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.