Economics and Human Biology 53 (2024) 101366

Contents lists available at ScienceDirect

Economics and Human Biology

journal homepage: www.elsevier.com/locate/ehb

Full length article

Lifetime costs of overweight and obesity in Italy

Vincenzo Atella [a,b,*], Federico Belotti [a,b], Matilde Giaccherini [b,c], Gerardo Medea [d], Antonio Nicolucci [e], Paolo Sbraccia [f], Andrea Piano Mortari [a,g]

[a] Department of Economics and Finance, Tor Vergata University of Rome, Italy
[b] CEIS Tor Vergata, Tor Vergata University of Rome, Italy
[c] CESifo, Germany
[d] Center for Outcomes Research and Clinical Epidemiology - CORESEARCH, Pescara, Italy
[e] Italian College of General Practitioners, Florence, Italy
[f] Department of Systems Medicine, Tor Vergata University of Rome, Rome, Italy
[g] Department Programming, Ministry of Health, Rome, Italy

ARTICLE INFO

Keywords:
Obesity
BMI
Lifetime expenditures
Primary care
Italy

ABSTRACT

We use longitudinal electronic clinical data on a large representative sample of the Italian population to estimate the lifetime profile costs of different BMI classes – normal weight, overweight, and obese (I, II, and III) – in a primary care setting. Our research reveals that obese patients generate the highest cost differential throughout their lives compared to normal weight patients. Moreover, we show that overweight individuals spend less than those with normal weight, primarily due to reduced expenditures beginning in early middle age. Our estimates could serve as a vital benchmark for policymakers looking to prioritize public interventions that address the obesity pandemic while considering the increasing obesity rates projected by the OECD until 2030.

1. Introduction

During the last decades, both industrialized and developing countries have seen significant increases in body weight and BMI due to unhealthy lifestyles, such as sedentary behavior and high-calorie food intake (Egger and Dixon, 2014; Popkin et al., 2006). The direct consequence of these changes has been a significant increase in long-term morbidity and chronic diseases, which then translated into a relevant public health problem (Atella et al., 2015; Ezzati et al., 2002; Haslam and James, 2005; Kortt et al., 1998).

Evidence from the International Association for the Study of Obesity (IASO) shows that the prevalence of obesity (BMI $\geq$ 30 kg/m^2) has exceeded 20% in numerous European countries (IASO, 2006). More recently, the OECD estimated a steady increase in obesity rates until at least 2030 (OECD, 2017). Exceptionally high obesity rates are projected in the United States, where in 2030, almost 50% of the population could be obese. Similar forecasts exist in Mexico (40%) and England (35%). Although at a slower pace, obesity is also growing in other countries. For example, in Italy, projections show that the prevalence of obesity in 2030 (13%) could double its 2000 level.

Obesity imposes high external costs, which are likely to cause a decrease in social welfare (Thompson et al., 1999; Yang and Hall, 2008; Finkelstein et al., 2010, 2008; Fallah-Fini et al., 2017; Cawley, 2015). According to a systematic review on the obesity burden in Europe, obesity has a significant economic impact in different European countries, with associated costs ranging from 0.09% to 0.61% of national gross domestic income (Müller-Riemenschneider et al., 2008). In a more recent review, von Lengerke and Krauth (2011) found that all but one of the 14 studies that compared obese to non-obese groups found higher costs with obesity, regardless of the type of cost components (total or specific) examined. In terms of obesity costs in subgroups, men, groups with a high SES defined by education, income, occupational status, or a composite index of these indicators, and obese groups with coexisting physical conditions had higher overall costs. Similar results have been obtained for the United States. In particular, Wolf and Colditz (1998) and Thompson et al. (1998) showed that obesity costs account for up to 10% of US healthcare spending.

Rising obesity rates have become a significant global health concern in recent years, and Italy, often admired for its healthy Mediterranean

* Corresponding author at: Department of Economics and Finance, Tor Vergata University of Rome, Italy.
E-mail address: atella@uniroma2.it (V. Atella).

[1] The project PASSI is an Italian public health surveillance system which collects information continuously on lifestyles and behavioral risk factors of the population aged 18-69 years living in Italy, with specific concern to the onset of chronic non-communicable diseases and compliance with disease prevention programs.

https://doi.org/10.1016/j.ehb.2024.101366
Received 7 April 2023; Received in revised form 16 November 2023; Accepted 1 February 2024
Available online 6 February 2024

diet, has not been spared from this challenge. Regarding Italy, a cross-sectional study carried out on a sample of people 18+ from ten general practices in the Ravenna local health unit found that, compared to normal weight patients, the average annual cost of drugs for overweight and obese patients increased by 86% and 153%, respectively (Degli Esposti et al., 2006). Similarly, Colao et al. (2017) studied the prevalence and associated costs of obesity in a population of more than 20,000 people in Italy. Their findings show that obese adults who did not receive adequate support for their obesity used more prescription drugs, were hospitalized longer, required more specialized outpatient care, and had significantly higher costs than normal weight adults. Additionally, the "*Progress by local health units towards a healthier Italy*" project (PASSI), implemented by the National Health Institute (ISS), reveals a national increasing trend over time until 2017, followed by a slight decrease until 2022 (rising from 10.3% in 2008 to 10.7% in 2017, and then dropping to 10.5% in 2022). It is worth noting that when we examine the variations across regions, we observe a rising trend in the northern regions, which increased from 9.3% in 2008 to 10.2% in 2022, reaching the national average. In contrast, the southern regions experienced a peak in 2017, and then a decreasing trend until 2022 (10.9%).[1]

The increasing prevalence of obesity poses severe risks to the nation's health, as it can lead to non-communicable diseases (NCDs) such as diabetes, heart disease, and some forms of cancer and strain the aging population's resources. This phenomenon, in turn, threatens the financial sustainability of the Italian National Health System (NHS).

Although informative, this literature is mainly based on cross-sectional analyses, providing policy makers with limited knowledge about how the burden of obesity disease starts and progresses throughout the life cycle. This limitation is particularly relevant given that obesity prevention initiatives are motivated by health and economic considerations. Hence, the possibility of correctly estimating the costs of obesity throughout a person's life is of paramount importance to provide policymakers with a correct evaluation of the amount of money that might be saved if successful anti-obesity measures are taken and health interventions are better prioritized.

Furthermore, estimating the lifetime medical expenditure of obesity is essential because it allows us to solve the ambiguity that arises when we consider that while obesity increases the risk of diseases and their treatment (leading to higher annual medical costs), it also shortens life expectancy for especially high levels of BMI (Stevens et al., 1999; Allison et al., 2003; Fontaine et al., 2003). Depending on whether the mortality or expenditure rate is higher, some levels of obesity can result in lower lifetime medical costs. However, for the US, Finkelstein et al. (2008) find that "even after taking differential survival probabilities into account, costs attributable to obesity are positive for all race and gender strata".[2] Moreover, the life cycle approach allows us to sort out another crucial issue from the standpoint of those countries, such as the US, where a private health system for workers coexists with a public pension system. In this case, it is important to understand how obesity-related health expenditure changes after retirement age.[3]

As noted in Schell et al. (2021), estimates of lifetime costs of obesity healthcare are difficult to obtain and vary widely, in part because available data are limited and the techniques underlying the cost models are opaque. The ideal estimate must consider several important factors, including age-related weight gain, different life expectancies, identifiability, and the choice of the cost model (Schell et al., 2021). Despite these difficulties, several articles have attempted to estimate the lifelong costs of obesity in the last two decades (Allison et al., 1999; Thompson et al., 1999; Tucker et al., 2006; Lakdawalla et al., 2005; Daviglus et al., 2004). Unfortunately, these works had some limitations problems. For example, contributions by Allison et al. (1999) and Thompson et al. (1999) are biased by the use of the attributable fraction technique, which includes only a few disorders and does not properly account for confounding and effect modification. Tucker et al. (2006) uses a semi-Markov model, based on some difficult-to-verify assumptions, and for specific BMI levels. As such, their results differ noticeably from those of previously published studies. Daviglus et al. (2004) focuses on expenditures from age 65 years to death using clinical data while the analysis carried out by Lakdawalla et al. (2005), in a dynamic microsimulation setting, is based on self-reported BMI measures and is centered on obesity status at age 70.

These estimation problems originate mainly from inadequate data to measure the phenomenon, forcing researchers to use different datasets to track changes in BMI and to rely on simplifying assumptions. As recognized by Schell et al. (2021), ideally, any longitudinal data set on individual BMI data trajectories should include information that "would be nationally representative, recent and cover subjects from early adolescence to their deaths". Given that a dataset with such unique characteristics are very difficult to be found, Schell et al. (2021) defined a list of criteria to determine whether a dataset deserves inclusion in a study with the aim of estimating an age-related weight gain curve despite its shortcomings. According to these authors, "the dataset should cover over 10 years of subjects' lives, be relatively recent, have sufficient follow-up and low attrition rates, have a short time between observations, objectively measure height and weight, and be nationally representative".

One of such dataset is represented by the Health Search/IQVIA Health LPD Longitudinal Patient Database (HS), which collects Electronic Clinical Records (ECR) for a large nationally representative sample (more than 1.5 million individuals) of Italian patients aged 14+, assisted by a group of 800 general practitioners (GP), with a time span ranging from 2004 to 2018. This data set presents all the features listed by Schell et al. (2021). In particular, it is nationally representative, covers up to 15 years of subjects' lives, which allows for sufficient follow-up, is collected yearly, has low attrition rates (only due to death or GP's repeal) and reports objectively measured BMI status and clinically assessed disease diagnosis. Additionally, in contrast, to claim data, the HS database includes all patients registered on physician lists (and not only those who demand health services), a key feature which reduces selection problems based on health status. These characteristics imply that individual information is not subject to recall error. Ultimately, as a primary care dataset, the cost analysis focuses on outpatient expenses, such as medications, diagnostic tests, and specialist consultations. Although some information may not be available (*e.g.* costs associated with hospitalization or long term care), the availability of primary care data offers a unique and valuable perspective compared to prior literature. First of all, GPs play the role of "gatekeepers" of the Italian NHS, hence evaluating costs related to primary care visits is crucial because primary care serves as the foundation of healthcare delivery for individuals with various BMI classes and health statuses, thus offering more accurate calculations of each person's healthcare expenditures. Secondly, the primary care setting is the most suitable to investigate the health status of the Italian population since it is not subject to selection bias. Thirdly, the costs associated with primary care visits represent a significant, yet frequently overlooked, aspect of healthcare expenses, particularly when analyzing medical costs by BMI class. Indeed, BMI impacts a wide population and influences overall health and healthcare spending, making it essential when discussing affordability and accessibility. Many chronic conditions like diabetes, hypertension, and asthma are effectively managed through regular primary care visits. Proper management of these illnesses can control symptoms, prevent

[2] They estimate the lifetime costs of obesity by combining medical costs and survival data for 20-year-old adults based on age, race, sex, and BMI.

[3] Employers and insurers are prone to overlook these costs when deciding how much to invest in measures to prevent obesity as long as obesity costs are age dependent. Furthermore, successful measures to reduce obesity can result in higher costs due to longer survival and, therefore, longer participation in public health programs while improving health and reducing costs related to obesity.

complications, and enhance patients' quality of life. Finally, primary care costs analysis may also signal the presence of existing health disparities. When primary care becomes financially inaccessible for individuals, they may forgo necessary medical attention, resulting in exacerbations of chronic conditions and potential emergency room visits at higher costs. These cost burdens could have disproportionately higher negative impact on marginalized or underserved communities, which already may face barriers to healthcare access leading to worse health outcomes and increased long-term healthcare costs.

From a statistical perspective, our BMI presents some missing data due to the so-called ascertainment bias, which occurs when individuals with missing data are likely different (*i.e.*, healthier) from those with complete information (Sedgwick, 2015). Consequently, we take advantage of the available information to recover the missing information. In particular, we estimate a sample selection model (Heckman, 1979) that allows us to test and correct any potential biases arising from non-random missing outcome measures. We assume that the sample selection generated by the missing BMI data is driven by observable characteristics, such as health status and age. Then, we build 14 overlapping cohorts for different BMI classes (normal weight, overweight, and obese (I, II, and III)) in order to estimate cohort specific costs. Finally, we combine the cohorts to obtain the lifetime (15-85) direct expenditure of a "representative" individual for each BMI class. It is worth noting that we allow for time-varying BMI, *i.e.* individuals can change their BMI class as they age.

Our findings show that obese patients represent the largest cost component throughout life, while the costs of overweight patients are lower than those of normal weight patients from around age 53 onward.

The article is structured as follows: Section 2 introduces the HS database, Section 3 outlines the empirical strategies employed to recover missing BMI information and derives lifetime estimates of BMI expenditure, Section 4 showcases our primary findings, Section 5 delves into the limitations of our empirical analysis and proposes some potential improvements, while Section 6 offers concluding remarks.

2. Data

The study is based on data from the Health Search/IQVIA Health LPD Longitudinal Patient Database (HS), an Italian general practice registry, which contains ECR data from patients aged 14+ registered with a group of 800 GPs, distributed in all Italian regions to be representative of the Italian population of GPs.[4] GPs voluntarily agree to collect clinical data and attend training courses for data entry (Cricelli et al., 2003; Filippi et al., 2005). Furthermore, to be considered for participation in epidemiological studies, all GPs recruited must meet standard quality criteria with respect to coding levels, prevalence of well-known diseases, mortality rates, and years of recording activity (Filippi et al., 2005).[5]

A key feature of the HS database is that it includes all patients registered in the GP lists, thus avoiding selection bias based on health status, which is a rather standard problem with claim data. Each Italian resident, regardless of citizenship, is enrolled with a GP who acts as a gatekeeper for the system. There is no private GP sector. Therefore, physician ECRs provide information on all individuals registered with a physician, regardless of their health status.[6]

The database contains patient demographic data (age, sex, province of residence) that are linked through an encrypted patient code with their medical records (diagnosis, prescribed tests, test results, and specialist visits), drug prescription information (medication name, date of prescription, and number of days of supply), self-reported hospital admissions, and date of death. Although GPs collect information daily, for this analysis, the information was aggregated at the year level from 2004 to 2018. Whenever variables were recorded multiple times within a year (*e.g.*, diagnostic test values, visits, BMI levels, drug prescriptions), we either averaged (*e.g.*, BMI) or summed (*e.g.*, visits, diagnostic tests) them at patient level, depending on the variable type. This resulted in a single observation per patient per year. In the original sample, the GP panel is balanced, while the patient panel is unbalanced due to events such as mortality, migration, or transitions from pediatricians to GPs and between GPs. As such, the original sample includes a total of 1,551,561 patients and 18,693,405 observations during the period 2004–2018. For our purposes, from this initial sample we selected patients in the age range [15–95], which reduced the sample to 1,547,530 patients and 18,461,302 observations.[7]

Unfortunately, during the 2004–2018 period, we observe only 686,173 individuals (44.3% of the selected sample) with at least one BMI measure. This characteristic of the sample can lead to an overestimation of health expenditures due to the so-called ascertainment bias, which occurs when individuals with missing data are likely different (*i.e.*, healthier) from those with complete information (Sedgwick, 2015). This selection bias arises mainly because when there are no medical reasons to collect these measurements, the GP is likely not to record them in a systematic way. For example, patients with cardiometabolic diseases or high BMI are more likely to have this information recorded by their GP. In our initial sample, 67% of subjects over 40 years of age and with one or more chronic conditions have at least one BMI record, while only 37% of those under 40 years of age and without important health problems have a BMI record. The descriptive statistics of the two samples are shown in Table 1 while Table 2 reports some descriptive statistics by BMI class for those for which we observe the BMI.[8]

This data reveals that individuals with missing BMI data tend to be healthier and younger: 89% of them have no chronic conditions. Furthermore, their average age is 48, compared to the average age

[4] The Italian National Healthcare System (NHS), known as *Sistema Sanitario Nazionale - SSN*, is a universal system funded by general taxation that provides healthcare services, mostly free of charge at the point of care, to all residents of Italy. Within this system, GPs play the role of 'gatekeepers' of the system, as they are responsible for prescribing specialist consultations, diagnostic tests, and drugs.

[5] In terms of coding levels and consistency in individual medical and clinical history, the HS database adheres to strict "up to standard" quality criteria (Cricelli et al., 2003; Filippi et al., 2005). More details on the degree to which the database represents the adult Italian population can be found at https://www.healthsearch.it/ (Official website of the HS database). The database complies with European Union guidelines on the use of medical data for research and has previously been shown to be a valid data source for scientific research (see Belotti et al., 2022; Mazzaglia et al., 2009; Atella and Kopinska, 2014; Atella and Conti, 2014; Atella and D'Amico, 2015; Atella et al., 2017, 2019, 2015, among others).

[6] Of note, the Italian National Institute of Statistics (ISTAT) uses these data to complement the information collected with the annual national health survey (ISTAT, 2012); furthermore, the Italian Drug Agency has used the Health Search database routinely as a source for the National Report on Drug Use in Italy since 2004 (OSMED, 2021, 2022). Furthermore, researchers have observed "a high degree of overlap between the population represented in the Health Search database and what is reported by ISTAT" (Bianchini et al., 2014).

[7] There are two reasons why we limit our sample to the age range of 15–95 in Italy. First, pediatricians are primarily responsible for treating people under the age of 14. In some cases, a pediatrician may see a patient up until they turn 16, after which transitioning to a general practitioner becomes mandatory. Thus, we chose to start our age range at 15 years old. This also allows us to construct our cohorts in five-year increments, beginning at age 15. Second, we cap our sample at age 95 due to dwindling numbers. It should be noted that in Section 3 we model life cycle expenditures up to the age of 85, as this is consistent with the average life expectancy in Italy.

[8] Concerning the BMI status, we have grouped the continuous variable into BMI classes as recommended by the WHO: normal weight (BMI between 18.5 and 24.99), overweight (BMI between 25 and 29.99), obesity class 1 (BMI between 30 and 34.99), obesity class 2 (BMI between 35 and 39.99), obesity class 3 (BMI ≥40).

Table 1
Descriptive statistics.

	Full sample	BMI observed	BMI missing
Age	50.48	53.392	47.736
	(19.395)	(18.496)	(19.819)
Gender (female)	0.53	0.542	0.519
	(0.499)	(0.498)	(0.500)
No chronic conditions	0.831	0.768	0.89
	(0.375)	(0.422)	(0.312)
Patients under 40 y.o.	0.34	0.264	0.411
	(0.474)	(0.441)	(0.492)
Coronary diseases	0.016	0.023	0.01
	(0.126)	(0.149)	(0.01)
Cerebrovascular diseases	0.034	0.047	0.022
	(0.181)	(0.212)	(0.145)
Heart failure	0.013	0.017	0.009
	(0.112)	(0.128)	(0.093)
Peptic ulcer disease	0.018	0.026	0.012
	(0.135)	(0.158)	(0.108)
Diabetes mellitus	0.075	0.114	0.038
	(0.263)	(0.318)	(0.191)
Solid tumor	0.01	0.012	0.008
	(0.098)	(0.107)	(0.087)
Depression	0.046	0.059	0.034
	(0.209)	(0.235)	(0.181)
Average direct costs (euro)	361.465	498	232.882
	(2344.483)	(3197.314)	(1000.527)
Average drugs costs (euro)	243.453	331.428	161
	(2309.225)	(3266.205)	(959.254)
Average outpatient costs (euro)	118.013	163	73.476
	(250.262)	(289.097)	(186.138)

Notes: Numbers refer to 1,547,530 individuals (18,461,302 obs.). BMI data refers to 686,173 individuals (1,751,551 obs.).

of 53 among those with BMI information. However, the full sample's average age is comparable to the Italian population's average age (45.2 in 2018 (Istat, 2023a)), considering we do not observe people under 15, who accounted for 13.2% of the population in 2018. The share of women is also in line with official statistics, with 53% in our sample and 52% for those aged 15 to 95 (Istat, 2023b). Regarding chronic conditions, diabetes prevalence in our sample is higher than that reported by ISTAT (7.5% versus 5.3% (Istat, 2017)), although ISTAT's prevalence estimate relies on self-reported health status while ours is based on diagnosed ICD9 codes. The heart failure prevalence in our full sample (1.3%) matches the figure reported in Buja et al. (2017).

Looking at the restricted sample of patients with observed BMI, as we can see from Table 2, the sample average age is 53.4 years. However, for underweight individuals, the average age drops to 38.7 years, increases to 57.1 for overweight individuals, and then decreases monotonically with obesity classes. In terms of gender, women are the most represented in the underweight, severely, and very severely obese classes. Finally, underweight problems occur mainly among people under 40 years of age (63%), while obesity is a more prevalent condition among people over 40 years of age. The evidence presented suggests that issues with being underweight primarily impact young women. Furthermore, while 4 out of 5 normal weight individuals do not have chronic diseases, only 40% of severely obese patients do not report chronic diseases.

Fig. 1 reports the average expenditures for drugs and outpatient services (diagnostics, laboratory tests, and specialists visits) by BMI class. We observe that expenditures are increasing in (adverse) BMI status: Obese I individuals face more than double drug-related expenditures and almost 50% higher outpatient costs compared to normal weight. It is worth recalling that the expenditures we consider are those associated with diagnosed diseases that are clinically connected with obesity. The approach of including a limited number of diseases has been criticized by Finkelstein et al. (2008) because it does not fully account for confounding and effect modification (Flegal et al., 2005). On the other end, as highlighted in the study by Guh et al. (2009), excess weight bears a significant clinical, social, and economic burden due to its association with a multitude of non-communicable diseases. These diseases include, but are not limited to, diabetes, cardiovascular conditions, respiratory ailments, certain forms of cancer, osteoarticular disorders, and depression. In our case, with a complete set of available ICD9 codes and diagnoses made by well-trained physicians, we believe that this approach can provide a more transparent and precise understanding of obesity costs. In particular, in this work we consider expenditures for patients who have been diagnosed with cerebrovascular disease (ICD9-CM: 431, 433[0,9], 434[0,9], 435, 436, V45.89); coronary diseases (ICD9-CM: V45[81,82], 410,412); Angina (ICD9-CM: 411, 413, 414), Heart failure (ICD9-CM: 428, 402[0,9], 404[11,93]) Diabetes mellitus (ICD9-CM: 250); Solid tumor: Colon (ICD9-CM: 153), pancreas (ICD9-CM: 157), stomach (ICD9-CM: 151), esophagus (ICD9-CM: 150), cholecyst (ICD9-CM: 156), liver (ICD9-CM: 155), intestinal (ICD9-CM: 152), rectum (ICD9-CM: 154) and chronic renal diseases (ICD9-CM: 585), Peptic ulcer disease (ICD9-CM: 531-534) and depression (ICD9-CM: 296).

3. The empirical strategy

In this section, we first describe how we recover missing information on BMI levels and then how we transform cross-sectional data into cohort data to reconstruct lifetime obesity costs.

3.1. Recovering missing information on BMI levels

As we have seen in Section 2, a major problem with our data is represented by the large share of missing BMI observations. To recover this missing information, useful for estimating cumulative lifetime expenditure related to obesity, several methods have been proposed in the literature (see among others Carpenter and Kenward, 2012; Reiter et al., 2006; De Silva et al., 2021). Taking advantage of the available information and assuming that the sample selection generated by the missing BMI values is driven by observable characteristics such as health status and age, we tried to reduce the impact of the ascertainment bias on our analysis. For our purposes, we use the sample selection model proposed by Heckman (1979) which provides a potentially useful tool, since it allows us to test and correct potential biases created by missing outcome measures that are not random. Our objective is not to impute the BMI value to each subject, which remains a very difficult task, but more simply to infer the BMI class to which each subject belongs. To this end, we exploit a sample selection model in which the response is the BMI level of each individual.

Our statistical model can be represented as follows:

$$y^*_{j,it} = \boldsymbol{x}_{j,it}\boldsymbol{\beta}_j + \epsilon_{j,it} \qquad j = 1,2 \tag{1}$$

$$y_{1,it} = I(y^*_{1,it} > 0) \tag{2}$$

$$y_{2,it} = y^*_{2,it} \qquad \text{if } y_{1,it} = 1 \tag{3}$$

where $y^*_{2,it}$ is the BMI of the subject i in year t, and $y^*_{1,it}$ identifies the selection process. Due to the missing data (selection mechanism), $y_{2,it}$ is observed only for the sub-sample of observation for which $y_{1,it} = 1$ (the selected sample). $\boldsymbol{x}_{1,it}$ and $\boldsymbol{x}_{2,it}$ are vectors of individual exogenous characteristics, such as age, sex, region of residence, fraction of individuals whose BMI was effectively recorded for the year t at the GP level and year fixed-effects. $\boldsymbol{x}_{2,it}$ also includes a vector of selected indicators of obesity-related diseases: cerebrovascular diseases, coronary diseases, heart failure, peptic ulcer disease, diabetes mellitus, and solid tumors (colon, pancreas, stomach, esophagus, cholecyst, liver, intestine, and rectum). Finally, $\epsilon_{1,it}$ and $\epsilon_{2,it}$ are idiosyncratic errors, and $\boldsymbol{\beta}_j$, $j = 1,2$, are the vectors of parameters to estimate.

Identification of vectors of unknown parameters $\boldsymbol{\beta}_j$, $j = 1,2$, requires an exclusion restriction in the selection equation (Eq. (2)), that is, $\boldsymbol{x}_{1,it}$ must contain at least one variable not contained in $\boldsymbol{x}_{2,it}$. We argue that this selection mechanism is driven by health status; thus,

Table 2
Descriptive statistics by BMI level (Observed BMI sample).

	Under	Normal	Over.	Obese I	Obese II	Obese III	Overall
Age	38.75	48.87	57.09	58.13	57.27	55.20	53.39
	(19.42)	(19.10)	(17.06)	(16.45)	(16.26)	(15.77)	(18.50)
Gender (female)	0.80	0.61	0.45	0.49	0.61	0.70	0.54
	(0.40)	(0.49)	(0.50)	(0.50)	(0.49)	(0.46)	(0.50)
No chronic conditions	0.90	0.84	0.74	0.67	0.63	0.60	0.77
	(0.30)	(0.37)	(0.44)	(0.47)	(0.48)	(0.49)	(0.42)
Under 40 y.o.	0.63	0.37	0.18	0.16	0.17	0.19	0.26
	(0.48)	(0.48)	(0.38)	(0.36)	(0.37)	(0.39)	(0.44)
Coronary diseases	0.00	0.01	0.03	0.03	0.03	0.02	0.02
	(0.07)	(0.12)	(0.17)	(0.17)	(0.16)	(0.14)	(0.15)
Cerebrovascular diseases	0.02	0.04	0.06	0.06	0.05	0.04	0.05
	(0.13)	(0.18)	(0.23)	(0.24)	(0.22)	(0.20)	(0.21)
Heart failure	0.01	0.01	0.02	0.03	0.03	0.04	0.02
	(0.08)	(0.10)	(0.13)	(0.16)	(0.18)	(0.19)	(0.13)
Peptic ulcer disease	0.01	0.02	0.03	0.03	0.03	0.02	0.03
	(0.11)	(0.14)	(0.17)	(0.17)	(0.16)	(0.15)	(0.16)
Diabetes mellitus	0.02	0.05	0.13	0.20	0.26	0.30	0.11
	(0.12)	(0.22)	(0.34)	(0.40)	(0.44)	(0.46)	(0.32)
Solid tumor	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(0.10)	(0.10)	(0.11)	(0.11)	(0.10)	(0.09)	(0.11)
Depression	0.05	0.05	0.06	0.07	0.07	0.08	0.06
	(0.22)	(0.23)	(0.23)	(0.25)	(0.26)	(0.27)	(0.23)
Average expenditures (€)	270.97	383.01	549.49	641.66	697.82	740.44	497.68
	(655.37)	(1213.81)	(5016.73)	(1403.38)	(1450.12)	(1107.40)	(3197.31)
Obs.	280,536	3,548,756	3,277,333	1,338,769	376,472	133,712	8,955,677
Patients	23,986	276,676	246,307	100,563	28,447	10,194	686,173

Notes: Numbers refer to mean and standard deviations (in parenthesis). Average expenditure refers to BMI-related outpatient (drugs, diagnostic tests, and specialist visits) per capita annual expenditure. All expenditures are defined in terms of 2018 constant prices.

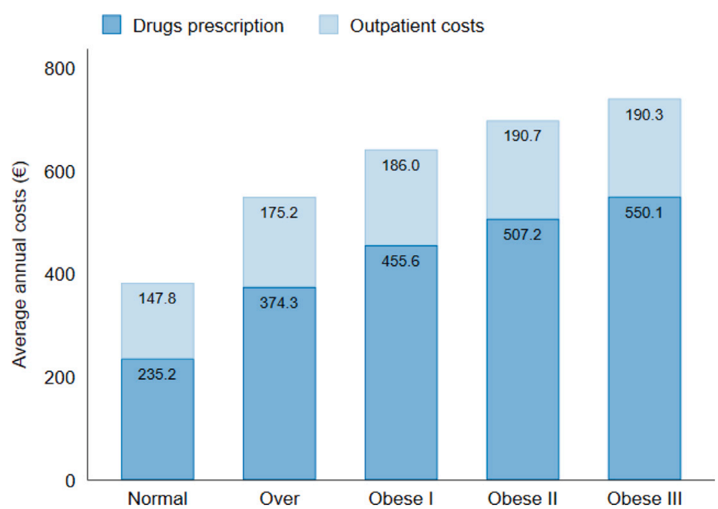


Fig. 1. Average obesity-related out-patient expenditures by type (2004–2018).
Notes: Bars represent the average obesity-related out-patient expenditures by drug prescription and other outpatient costs by BMI class (sample of 686,173 patients).

we consider as exclusion restriction two binary indicators such as being healthy and under 40 years of age.

We estimate the model (1)–(3) by pooling the unbalanced panel of 1,54l7,530 individuals aged [15 – 95] and observed over the 15-year period, which allows exploiting all available information. We run the same model using two versions of the BMI variable: in the first case, we use the available BMI observations, while in the second case, we exploited an "interpolated version" of the BMI.[9] Table A2 in the Appendix reports the coefficients of the two specifications. The main equation shows the relationship between BMI and medical conditions, while the selection equation shows the relationship between the probability that an individual will have their BMI recorded and medical conditions. Our estimates show that there is a statistically significant positive relationship between heart failure and BMI, as well as between diabetes mellitus and BMI. Furthermore, the likelihood that a subject will have his BMI recorded is negatively related to having no chronic diseases and being under 40 years old. The Mills ratio, which is a measure of the strength of the relationship between the selection equation and the main equation, is also shown to be statistically significant. Furthermore, we computed the marginal effects of the exclusion restrictions included in the selection equation and found both to be negative and strongly significant. In general, these results suggest that the sample selection model used in this analysis effectively controls potential biases in the data.

Before proceeding with the BMI class imputation, it is worth noting that individuals with accessible BMI data display an average transition

[9] For individuals with more than one BMI entry over our observation period, we linearly interpolate the values between the two observations, but we do not extrapolate them.

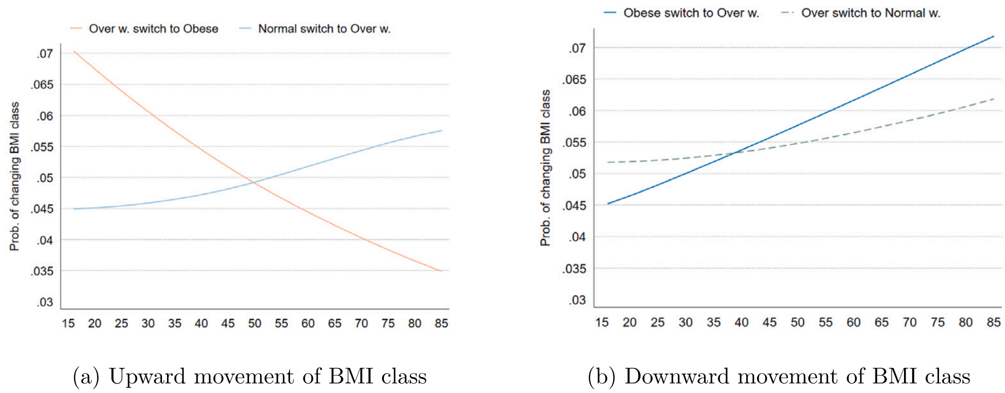


Fig. 2. Probability of changing BMI class (Interpolated BMI before imputation).

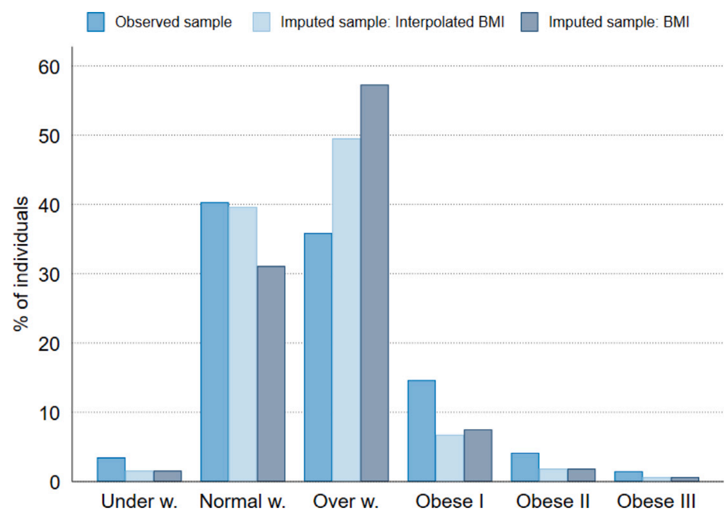


Fig. 3. BMI class before and after imputation
Notes: The figure shows the BMI class distribution for the observed sample (686,173 patients), the sample imputed with the entry (first year) BMI level and the interpolated BMI (1,547,530 patients for the last two samples).

probability of about 5% for switching between BMI classes, as illustrated in Fig. 2.[10] In panel (a), we can see a rising likelihood of weight gain with age for individuals who initially fall under the normal weight category. Conversely, there is a declining probability of transitioning to obesity for those who start as overweight. Panel (b), on the other hand, illustrates the opposite scenario for individuals who begin as obese or overweight, revealing a high probability of losing weight with age. This finding highlights the importance of integrating a time-varying BMI imputation.

Hence, we use the estimated parameters to impute $y_{2,it}$ when $y_{1,it} = 0$ assigning the individual BMI values to allow for time-varying BMI. Then, we categorize individuals into BMI classes that may change as they age.

It is crucial to highlight that when we have access to observed BMI data, we classify individuals based on the actual measurements rather than relying on predicted values.

As a check for our imputation strategy, in Fig. 3 we show the prevalence of BMI classes for the observed BMI sample and for the imputed BMI outcomes.[11] Regarding the distribution of observed BMIs, using the interpolated BMI as an outcome returns a similar prevalence of normal weight individuals and a higher prevalence of overweight individuals, while the prevalence of obese individuals is reduced. This result seems to point towards the interpolated BMI as our preferred outcome for the Heckman selection model, since if we use the BMI, we obtain a lower prevalence of normal weight individuals and a higher one of overweight. Although we expected a higher share of normal weight individuals after the imputation procedure, the results are quite satisfactory, as all prevalence values in the obese class are reduced compared to the original data.

Finally, in Table 3 we report the descriptive statistics of our sample by BMI class. Compared to Table 2, the results are not very different. In general, we can find differences only in terms of age (50.5 imputed *vs.* 53.4 observed), absence of chronic conditions (83% *vs.* 77%) and the number of people under 40 years of age (34% *vs.* 26%). Finally, the total expenditure per capita is lower in the imputed sample (361.5 euros) compared to the original sample (497.7 euros). This evidence suggests that the imputation process has produced a rebalancing of the BMI measurement versus those who are usually not recorded by the GPs because their health status is rather good. Following the imputation procedure, we have a younger sample, with a higher number of people without chronic diseases who spend less. Furthermore, among individuals with excess weight, those in the "Obesity III" class have the lowest mean age, while those in the "Obesity I" class tend to be older. The prevalence of women increases with increasing BMI from 48% among overweight individuals to 71% among severely obese

[10] We estimate four logit models to analyze the probability of transitioning between BMI classes. We classify individuals into BMI classes using the interpolated BMI data observed prior to imputation. The probabilities of changing BMI classes over time are presented in Table A4.

[11] Figure A2 in the Appendix reports Kernel density estimates of the observed, and imputed BMI values.

Table 3
Descriptive statistics by BMI class (after imputation).

	Under	Normal	Overw.	Obese I	Obese II	Obese III	Overall
Age	38.749	41.685	54.739	58.386	57.272	55.204	50.480
	(19.424)	(21.479)	(16.315)	(16.260)	(16.264)	(15.771)	(19.395)
Gender (female)	0.801	0.607	0.477	0.481	0.608	0.699	0.530
	(0.400)	(0.488)	(0.499)	(0.500)	(0.488)	(0.459)	(0.499)
Under 40 y.o.	0.627	0.586	0.217	0.149	0.166	0.189	0.340
	(0.484)	(0.493)	(0.412)	(0.357)	(0.372)	(0.392)	(0.474)
No chronic conditions	0.898	0.889	0.830	0.648	0.630	0.604	0.831
	(0.303)	(0.314)	(0.376)	(0.478)	(0.483)	(0.489)	(0.375)
Coronary diseases	0.005	0.009	0.018	0.033	0.028	0.019	0.016
	(0.068)	(0.096)	(0.132)	(0.179)	(0.165)	(0.135)	(0.126)
Cerebrovascular diseases	0.017	0.026	0.035	0.061	0.053	0.041	0.034
	(0.131)	(0.159)	(0.184)	(0.239)	(0.225)	(0.198)	(0.181)
Heart failure	0.007	0.007	0.013	0.030	0.032	0.038	0.013
	(0.082)	(0.083)	(0.113)	(0.169)	(0.177)	(0.192)	(0.112)
Peptic ulcer disease	0.012	0.014	0.019	0.031	0.027	0.024	0.018
	(0.111)	(0.118)	(0.137)	(0.172)	(0.162)	(0.153)	(0.135)
iabetes mellitus	0.015	0.030	0.072	0.233	0.260	0.298	0.075
	(0.122)	(0.172)	(0.258)	(0.423)	(0.438)	(0.458)	(0.263)
Solid tumor	0.010	0.010	0.009	0.012	0.010	0.009	0.010
	(0.099)	(0.099)	(0.095)	(0.111)	(0.102)	(0.094)	(0.098)
Depression	0.053	0.038	0.046	0.065	0.074	0.079	0.046
	(0.225)	(0.192)	(0.210)	(0.246)	(0.261)	(0.270)	(0.209)
Average expenditures (€)	270.971	272.170	360.582	652.017	697.822	740.439	361.465
	(655.369)	(1079.959)	(2985.061)	(1405.701)	(1450.119)	(1107.399)	(2344.483)

Notes: Numbers refer to mean and standard deviations (in parenthesis). Average expenditure refers to out-patient obesity-related per capita annual expenditure.

individuals, which is also compatible with the longer life expectancy of women. Table 3 also shows that individuals with higher levels of obesity tend to have higher average expenditures for outpatient medical treatment. This is consistent with previous research in the field of health economics, which has shown that obesity is associated with an increased risk of a variety of chronic health conditions and that treatment of these conditions can be costly (Kortt et al., 1998; Atella et al., 2015). For example, obesity is an important risk factor for the development of type 2 diabetes, which is a chronic disease that requires ongoing medical care and can result in significant medical costs (Anderson et al., 2003). Obesity is also associated with an increased risk of heart disease and stroke, two conditions that are both expensive to treat and leading causes of disability and death (DeCaria et al., 2012).

3.2. The cohort approach and the representative individual construction

To mitigate attrition issues stemming from mortality, migration between General Practitioners (GPs), or transitioning from pediatricians to GPs, and to generate a dataset that can coherently estimate lifetime obesity costs, we narrowed our sample to a balanced panel of individuals present from 2004 to 2018. This process yielded a sample that is balanced in terms of GPs and patients, comprising 856,620 patients and 12,849,300 observations. Descriptive statistics of this subsample are reported in Table A1 together with a test that compares averages of selected variables across the two sample. It is worth noting that, due to the large sample size, although the differences in averages are statistically significant for nearly all comparisons, the magnitudes of these differences are very small and can be considered insignificant from an epidemiological point of view. Furthermore, we disregarded the underweight class given the additional and specific medical care costs associated with being in this group.

We organized the data using a cohort approach, restricting our sample only to patients who, in 2004, were 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, and 80 years old and followed for 15 years. This setting produced 14 overlapping cohorts and limited further the number of individuals to 176,927 (due to the exclusion of all patients who, in 2004, were not aged 20, 25, 30, etc.).[12] As will become clear later, the overlapping nature of the cohorts is important for the construction of lifetime cost profiles. Table 4 shows the descriptive statistics of this new sample by BMI class and overall.

The final step consists of combining the cohorts to obtain the lifetime (15-85) direct expenditure on obesity of an individual representative of each BMI class (*e.g.*, an individual who is consistently obese from 15 to 85 years).[13] First, within each cohort, each patient starts with zero obesity-related expenditure in January 1st 2004 and accumulates costs until December 31st, 2018. This implies that, on January 1st 2004, a patient aged 15 and belonging to the [15–29] cohort has zero expenditures as any other patient in any other cohort. However, being interested in lifetime expenditure accumulation, we need to move our research interest from "flows" to "stocks" variables. This implies letting each cohort after the first ([15–29]) to accumulate not from zero, but from the end point of the previous cohort. By proceeding sequentially, each cohort after the first starts accumulating from the previous stock. In a more formal way, this is the process we follow.

Let $i = 1, \ldots, N$ denotes the individual, $t = 2004, \ldots, 2018$ the time, $b = 1, \ldots, 5$ the BMI class, $j = 1, \ldots, 14$ the cohort and $a = l_j, \ldots, u_j$ the lower and upper age bounds of cohort j. Then the cumulative average expenditure for each BMI class b and cohort j can be represented as:

$$\Phi_j^b = \sum_{a=l_j}^{u_j} \sum_{t=2004}^{2018} E_{j,a,t}^b \tag{4}$$

where $E_{j,a,t}^b = \frac{1}{N}\sum_{i=1}^{N} E_{i,j,a,t}^b$ is the direct average expenditure at time t, for age a of BMI class b. Notice that individuals can change BMI class over time.

Therefore, the cumulative "lifetime" (15-85) direct expenditure for BMI class b is:

$$\Phi^b = \sum_{j=1}^{14} \Phi_j^b \tag{5}$$

[12] The cohort age classes are: [15–29]; [20–34]; [25–39]; [30–44]; [35–49]; [40–54]; [45–59]; [50–64]; [55–69]; [60–74]; [65–79]; [70–84]; [75–89]; [80–94].

Figure A1 in Appendix 1 shows the criteria used to select the sample, with the relative sample sizes.

[13] We set the maximum age at 85, which is the average life expectancy in Italy (Raleigh, 2019).

Table 4
Cohort level data means and standard deviations by BMI class.

	Normal	Over	Obese I	Obese II	Obese III	Overall
Age	44.288	54.287	58.389	57.356	54.493	51.479
	(19.877)	(15.512)	(15.376)	(15.162)	(14.901)	(17.740)
Gender (female)	0.616	0.479	0.484	0.615	0.713	0.531
	(0.486)	(0.500)	(0.500)	(0.487)	(0.452)	(0.499)
No chronic conditions	0.873	0.833	0.652	0.633	0.610	0.825
	(0.333)	(0.373)	(0.476)	(0.482)	(0.488)	(0.380)
Under 40 y.o.	0.511	0.209	0.137	0.146	0.188	0.298
	(0.500)	(0.407)	(0.343)	(0.353)	(0.391)	(0.458)
Coronary diseases	0.010	0.017	0.031	0.028	0.014	0.016
	(0.100)	(0.130)	(0.173)	(0.164)	(0.119)	(0.126)
Cerebrovascular diseases	0.028	0.034	0.058	0.055	0.042	0.034
	(0.164)	(0.181)	(0.233)	(0.228)	(0.200)	(0.182)
Heart failure	0.006	0.010	0.023	0.027	0.029	0.010
	(0.076)	(0.097)	(0.150)	(0.163)	(0.168)	(0.100)
Peptic ulcer disease	0.016	0.020	0.032	0.026	0.025	0.020
	(0.127)	(0.140)	(0.175)	(0.159)	(0.155)	(0.140)
Diabetes mellitus	0.035	0.068	0.228	0.255	0.289	0.076
	(0.184)	(0.251)	(0.419)	(0.436)	(0.453)	(0.266)
Solid tumor	0.010	0.008	0.012	0.011	0.009	0.009
	(0.099)	(0.088)	(0.109)	(0.103)	(0.097)	(0.094)
Depression	0.047	0.049	0.067	0.074	0.092	0.051
	(0.212)	(0.215)	(0.249)	(0.261)	(0.289)	(0.219)
Average expenditures (€)	301.660	353.243	633.762	677.564	701.720	370.087
	(715.448)	(1385.165)	(983.777)	(987.617)	(1037.512)	(1171.461)

Notes: Numbers refer to mean and standard deviations (in parenthesis). Average expenditure refers to out-patient obesity-related per capita annual expenditure.

In addition, by further aggregating individual costs across different individual characteristics, we can obtain other meaningful profiles, for example, by gender and geographical areas.

4. Empirical results

Fig. 4, plots the evolution over time of the cumulative average expenditure of the individuals belonging to cohort j, *i.e.* $E^b_{j,t} = \sum_{a=l_j}^{u_j} E^b_{j,a,t}$. Each panel in the figure compares the average expenditure of normal weight individuals with that of other BMI classes (*e.g.*, in Fig. 4(a), the first solid and dashed lines represent the evolution over time for normal and overweight individuals aged 15 in 2004). Interestingly, compared to individuals with normal weight, the cumulative average expenditure tends to increase with each BMI category, except for overweight people who, beginning from the cohort [35-49], exhibit a lower cumulative average expenditure (Fig. 4(a)).

In Fig. 5, we display the same data using an alternative graphical representation. As observed, as individuals age, the impact of negative health shocks on expenditures becomes more pronounced. However, their magnitude increases significantly and consistently with higher BMI classes.[14] One way to gain a better understanding of the interaction between the two phenomena is to focus on Obese III individuals (as shown in Fig. 4(d)) and compare their average expenditure over a period of six years (35-40) to that of a normal weight individual over a longer period of 15 years ([25–39]). By doing this, we can observe that the two values are comparable. Specifically, Obese III individuals spent 1,855 euros in six years, while normal weight individuals accumulated 1,880 euros over 15 years (see Fig. 5).

Finally, using Eq. (4), we calculated the cumulative lifetime average outpatient expenditure profiles for the normal weight, overweight, and the three obese classes, which are shown in Fig. 6.

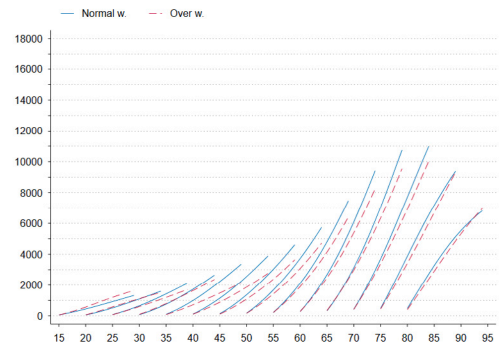
As expected, the difference across BMI classes is lowest during the teenage years and then increases continuously thereafter. Examining the specific pattern presented in Fig. 7, we can observe in panel (a) that starting from early middle age (39 y.o.), the expenses linked to being overweight are lower than those associated with the normal weight category, although statistically different only starting at the age of 53. This distinction continues to widen over time, a trend that has been previously noted in various contexts. (von Lengerke et al., 2006; Finkelstein et al., 2008; Tsai et al., 2011). As expected, cumulative average expenditures indicate that obesity classes are the largest cost component at all time points.[15]

Table 5 presents the cumulative average outpatient expenditure at 85 years by sex, geographic area, and health status. As expected, and consistently with international studies (Cawley et al., 2015), our data reveals a positive correlation between increasing BMI and lifetime expenditures. Individuals classified as Obese III exhibit the highest lifetime expenditures, reaching 43,950.97 euros. This underscores the escalating economic burden associated with severe obesity. When examining gender differences, and contrary to general trends observed in other studies (*e.g.* von Lengerke and Krauth (2011)) our data reveals that, on average, women exhibit higher lifetime expenditures than men across all BMI classes. This surprising finding may warrant further investigation into the specific healthcare utilization patterns and cost drivers for women, potentially indicating gender-specific healthcare needs. Disparities in lifetime expenditures across Northern, Central, and Southern areas are in line with the finding from Bruzzi et al. (2022) and highlight potential regional variations in healthcare utilization and costs. For instance, individuals in the Central area exhibit the highest expenditures, particularly in the Obese III category, while the Northern Area has the lowest expenditures across all BMI classes. These differences could reflect differences in healthcare practices, infrastructure, or disease prevalence, highlighting the need for region-specific health policies. Lastly, the table underscores the substantial economic impact of chronic conditions. Individuals with two or more chronic conditions demonstrate significantly higher lifetime expenditures. The pronounced increase in expenditures with the number of chronic conditions mirrors the findings of Lehnert et al. (2011).

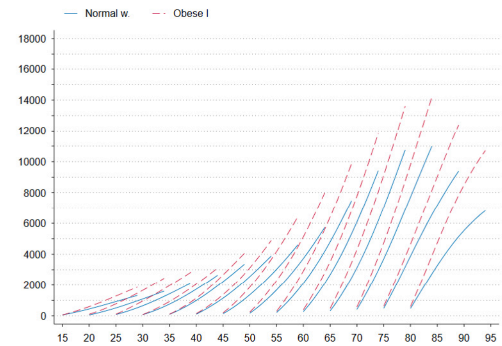
Comparing these findings with international studies, such as those from the United States (Finkelstein et al., 2008), Europe (Müller-Riemenschneider et al., 2008) and wordwide (Withrow and Alter, 2011) underscores the universality of the economic impact of obesity. Our analysis contributes to the global discourse on obesity-related

[14] Due to the small sample size for the cohort [15–29], the per capita annual expenditure for those who are class II obese at age 15 is lower than it is for those who are normal weight.

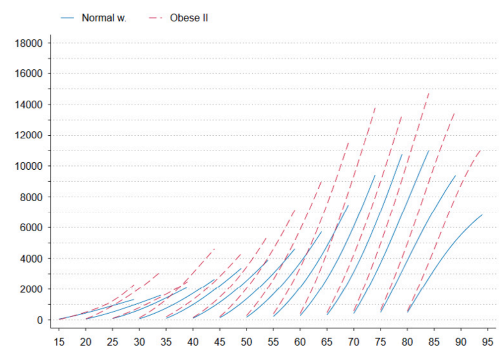
[15] Table A5 displays the results of mean comparison tests on the lifetime cumulative average obesity-related expenditure by BMI class.



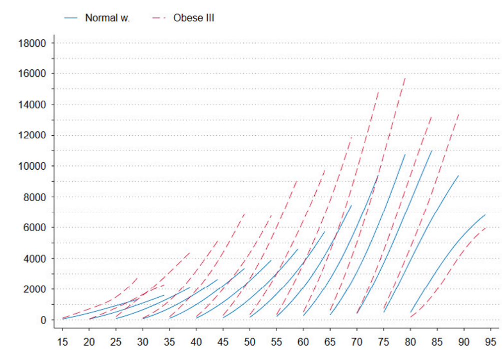
(a) Normal VS Over



(b) Normal VS Obese I



(c) Normal VS Obese II



(d) Normal VS Obese III

Fig. 4. Cumulative average BMI class expenditure by cohort
Notes : Age in 2004 and BMI at entry are used to identify cohorts. The figures correspond to 176,927 profiles that were tracked from 2004 to 2018 over a 15-year period. Direct expenses include outpatient expenses and drug costs.

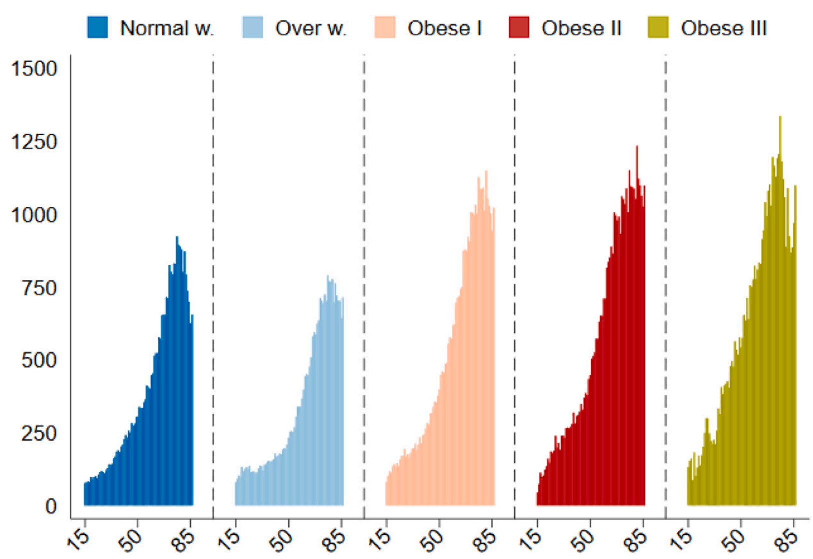


Fig. 5. Average BMI class expenditure by age .

healthcare costs, reinforcing the need for coordinated international efforts to mitigate the economic burden of obesity.

Figure A3 reports the cumulative average expenditure by cohort, BMI classes and type of expenditure. According to this evidence the main driver of the increase in expenditure for obese individuals, with respect to normal weight and over weight ones, is drugs prescription while other outpatients costs remain fairly similar for the same cohort across BMI classes.

Lastly, to assess the importance of differences in life expectancy (LE) between normal weight individuals and those at different classes of obesity we computed lifetime expenditure profiles based on the estimated LE by gender and BMI class in Bhaskaran et al. (2018). In particular, we applied the following reduction in LE, relative to an average LE for normal weight male of 82 y.o. and female of 85 y.o.:

- Overweight: Male: −1 y.; Female: −2 y.

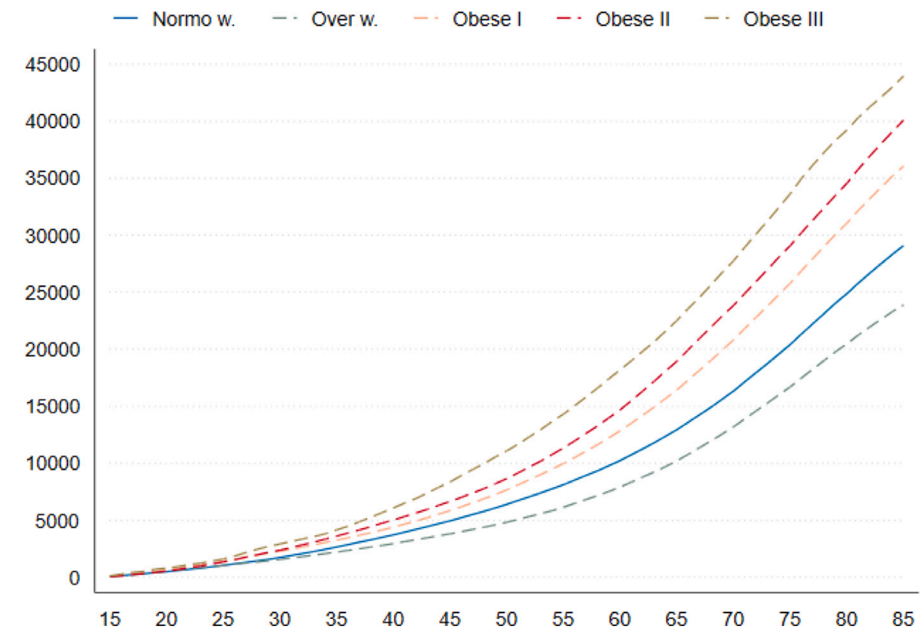
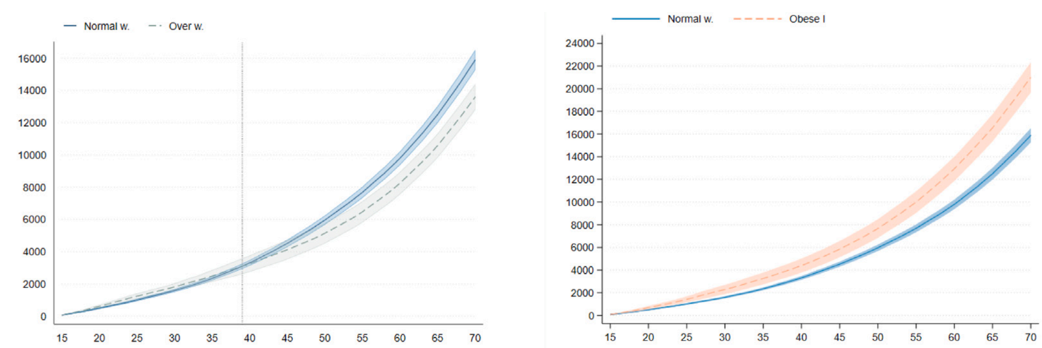


Fig. 6. Lifetime cumulative average BMI class expenditure .



(a) Normal *vs* over w.

(b) Normal *vs* obese I

Fig. 7. Details of the lifetime cumulative average BMI class expenditure.

Table 5
Lifetime average BMI class expenditure by sociodemographic status at age 85.

	Normal	Over	Obese I	Obese II	Obese III
All	28 076.57	24 446.74	36 609.27	40 092.39	43 950.97
	(2320.285)	(2045.749)	(2965.973)	(3114.747)	(3038.217)
Male	28 369.16	24 478.35	36 772.33	38 906.2	43 895.12
	(2671.374)	(2250.56)	(3317.883)	(3432.557)	(3057.644)
Female	28 057.87	24 954.16	37 345.14	41 741.99	43 944.97
	(2147.108)	(1836.169)	(2657.611)	(2902.463)	(3088.528)
Northern area	26 275.94	23 574.91	35 246.23	40 242.78	39 749.91
	(2140.119)	(2025.103)	(2902.709)	(3008.017)	(2910.198)
Central area	28 557.38	25 630.44	35 906.14	38 534.13	49 826.62
	(2383.268)	(2032.378)	(3129.097)	(3273.107)	(3650.961)
Southern area	31 457.48	25 107.43	37 980.8	40 598.27	45 328.45
	(2684.344)	(2074.614)	(2993.145)	(3193.176)	(3230.491)
w/o chronic conditions	23 331.95	19 318.28	27 906.1	30 322.55	31 460.38
	(1753.867)	(1428.282)	(2094.128)	(2129.613)	(2079.828)
One chronic condition	42 921.9	42 460.58	54 128.88	51 699.55	59 029.84
	(2109.343)	(2181.046)	(2384.528)	(3139.026)	(2986.963)
≥2 chronic conditions	68 938.67	68 388.12	75 795.52	68 251.87	79 208.6
	(3642.516)	(3868.414)	(4902.929)	(5571.668)	(5971.499)

Notes: The Northern area includes of eight regions: Aosta Valley. Piedmont. Liguria. Lombardy. Emilia-Romagna. Veneto. Friuli-Venezia Giulia. and Trentino-Alto Adige. The Central area includes four regions: Lazio. Umbria. Marche. and Tuscany. Finally. the Southern area includes eight regions: Abruzzo. Apulia. Basilicata. Calabria. Campania. Molise. Sicily. and Sardinia. Values are expressed in 2018 constant €. Standard errors in parenthesis.

- Obese I: Male: −3 y.; Female: −2 y.
- Obese II: Male: −6 y.; Female: −5 y.
- Obese III: Male: −9 y.; Female: −8 y.

Fig. 8 shows the result of this exercise and signals how the ranking in terms of expenditure is not significantly sensitive to differential life expectancy by BMI class, although the expenditure levels are lower (about 10,000 euro less for the Obesity III class).

5. Limitations

In conclusion, it is important to acknowledge certain limitations in our analysis.

Firstly, our study heavily relies on imputing the BMI class for individuals without recorded BMI values and employing linear interpolation for observed BMI levels. While our empirical strategy addresses this issue, it may still pose a limitation.

Secondly, the decision to use a balanced panel for cohort construction results in a nearly 50% reduction in sample size. Despite the minor differences between the two samples (as shown in Table A1), it is crucial to note that they represent distinct populations. Additionally, the balanced panel does not account for deceased individuals and their associated peak expenditures in the last year of life (Atella and Conti, 2014).

Thirdly, opting for a "representative individual" simplifies our cohort analysis but sacrifices heterogeneity.

Fourthly, our study is related to the primary care setting, which lacks information on hospitalization costs and long-term healthcare expenses. Consequently, our estimates for BMI-related costs could be interpreted as a lower bound. In fact, by leaving out these important cost components, our analysis might not fully capture the total economic impact of BMI-related problems.

Fifthly, it is important to note that our study does not include age-specific cost discounting. This decision is based on the nature and focus of our research. Time-based discount factors are more relevant in situations where individuals strategically optimize their health expenditures over time, but we view our approach as more of an "accounting exercise". Additionally, in a public health context, expenditure value is considered uniform regardless of the spender's age. Moreover, within the Italian NHS setting, it is worth mentioning that individuals do not directly bear the costs; for example, the entire expense for someone with a chronic condition is covered by the NHS.

Sixthly, while this approach enables a lifetime estimation of expenses associated with overweight and obesity, its primary objective is to investigate the relationship between obesity and healthcare costs rather than proving causality. This limitation stems from the lack of a reliable instrument, as pointed out by Cawley and Meyerhoefer (2012).

In conclusion, it should be noted that achieving cost-saving goals may become increasingly difficult due to the introduction of expensive treatments. In such instances, high-cost treatments may be "cost-effective" rather than "cost-saving". This technological aspect positively impacts our lifetime cost estimates, which we have not been able to isolate. However, in the specific case of Semaglutide, a recent high-cost drug for weight reduction, this does not pose a concern in our context since our observation period ends before its commercialization in Italy.

6. Conclusion

The OECD predicts that the number of obese people will continue to increase until at least 2030. This will affect the health of the population because obesity is linked to many health problems, such as coronary heart disease, chronic kidney failure, different types of cancer, sleep apnea, gallbladder disease, Type 2 diabetes, and more. Furthermore, as evidenced by numerous studies on the impact of obesity on healthcare spending, obesity is one of the main drivers of rising healthcare costs, primarily as a result of increased spending on chronic diseases caused by obesity.

By estimating lifetime expenditure associated with different levels of BMI at the population level, this study adds new evidence to the existing literature. In particular, using primary care data from the Health Search/IQVIA Health LPD Longitudinal Patient Database (HS), a large Italian general practice registry, distributed across all Italian regions, we study how patients belonging to different BMI classes (normal weight, overweight, and obese (I, II, and III)) generate differential lifetime cost profiles for the Italian national healthcare system in a primary care setting.

According to our research, obese individuals generate the highest expenditure throughout the life cycle, the main driver of the increase in expenditure is drugs prescription and women exhibit higher lifetime expenditures than men across all BMI classes. These findings have several policy implications, aimed at addressing the economic impact of obesity and overweight which requires a multifaceted approach that involves public health initiatives, healthcare system improvements, and targeted interventions.

In particular, our results suggest to implement and strengthen preventive health policies that focus on obesity prevention and health promotion from an early age. School-based programs, community interventions, and public awareness campaigns can play a crucial role in encouraging healthy lifestyles. In this sense, it could be useful to explore the implementation of financial incentives or health promotion programs to encourage healthier behaviors and weight management. These could include subsidized gym memberships, wellness programs at workplaces, or policies that promote access to healthy food options.

A specific role may be played also by those policies that can address gender-specific healthcare interventions to address the gender disparities observed in lifetime expenditures. Understanding the unique healthcare needs of men and women can guide the development of gender-specific programs and services, promoting more effective and equitable health outcomes.

Another policy relevant suggestion comes from the evidence of regional differences in healthcare expenditures associated with BMI classes. In this case it could be worth to address the problem by developing region-specific health plans. This may involve allocating resources based on the specific healthcare needs and challenges faced by different areas, promoting better healthcare access and outcomes.

Finally, policies should aim to strengthen chronic disease management programs, especially for individuals with multiple chronic conditions. Integrated care models, care coordination, and patient education can contribute to better management of chronic conditions, potentially reducing the long-term economic burden on the healthcare system.

By integrating these policy consequences into a comprehensive strategy, policymakers can work towards mitigating the economic impact of obesity and over weight, improving public health outcomes, and creating a more sustainable and equitable healthcare system.

Funding

This work was supported by Novo Nordisk Italy (Contract N. 8000347757), which had no role in the study design, conduct of the study, collection, management, analysis and interpretation of the data, or the preparation and review of the manuscript. All authors read and approved the final manuscript.

Ethics approval

The study did not require ethics approval.

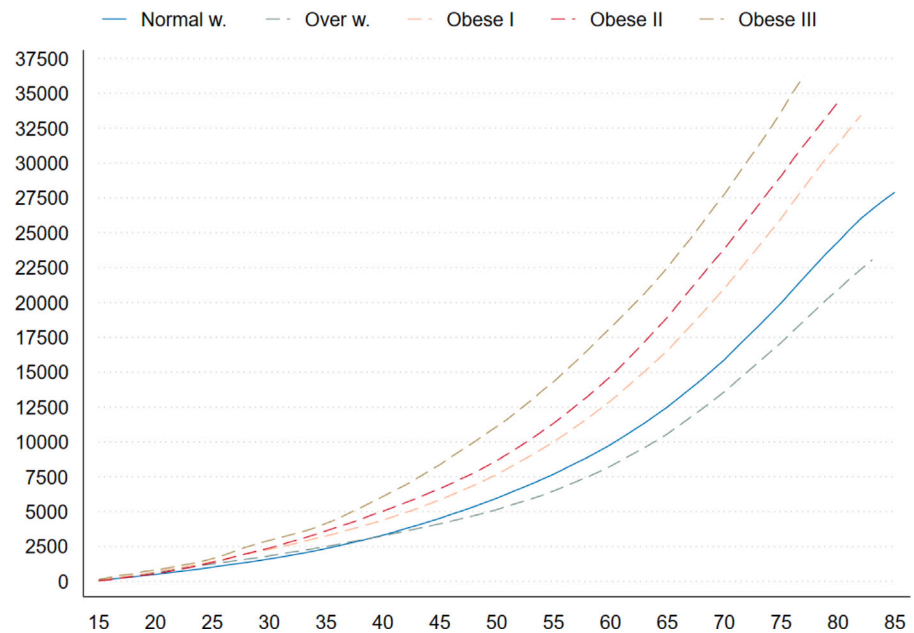


Fig. 8. *Scenario*: Heterogeneous Life Expectancy by BMI class.
Notes: This figure corresponds to 176,927 profiles that were tracked from 2004 to 2018. Average expenditures refer to per capita annual expenditures and include outpatient expenses and drug costs.

CRediT authorship contribution statement

Vincenzo Atella: Contributed to the study conception and design, Material preparation, Data collection, Analysis. **Federico Belotti:** Contributed to the study conception and design, Material preparation, Data collection, Analysis. **Matilde Giaccherini:** Contributed to the study conception and design, Material preparation, Data collection, Analysis. **Gerardo Medea:** Writing and revision of the manuscript. **Antonio Nicolucci:** Writing and revision of the manuscript. **Paolo Sbraccia:** Writing and revision of the manuscript. **Andrea Piano Mortari:** Contributed to the study conception and design, Material preparation, Data collection, Analysis.

Declaration of competing interest

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

Data availability

The data that has been used is confidential.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ehb.2024.101366.

References

Allison, D.B., Westfall, A.O., Redden, D.T., Wang, C., Fontaine, K.R., 2003. Methods of estimating years of life lost due to obesity. JAMA 289 (22), 2941–2942.
Allison, D.B., Zannolli, R., Narayan, K., 1999. The direct health care costs of obesity in the United States. Am J Public Health 89 (8), 1194–1199.
Anderson, J.W., Kendall, C.W., Jenkins, D.J., 2003. Importance of weight management in type 2 diabetes: Review with meta-analysis of clinical studies. J. Am. Coll. Nutr. 22 (5), 331–339, PMID: 14559925.
Atella, V., Belotti, F., Bojke, C., Castelli, A., Grasic, K., Kopinska, J., Mortari, A.P., Street, A., 2019. How health policy shapes healthcare sector productivity? Evidence from Italy and UK. Health Policy 123 (1), 27–36.
Atella, V., Belotti, F., Depalo, D., 2017. Drug therapy adherence and health outcomes in the presence of physician and patient unobserved heterogeneity. Health Econ. 26, 106–126.
Atella, V., Conti, V., 2014. The effect of age and time to death on primary care costs: The Italian experience. Soc. Sci. Med. 114, 10–17.
Atella, V., D'Amico, F., 2015. Who is responsible for your health: is it you, your doctor or the new technologies? Eur. J. Health Econ. 16, 835–846.
Atella, V., Kopinska, J.A., 2014. The impact of cost-sharing schemes on drug compliance in Italy: evidence based on quantile regression. Int. J. Public Health 59 (2), 329–339.
Atella, V., Kopinska, J., Medea, G., Belotti, F., Tosti, V., Piano Mortari, A., Cricelli, C., Fontana, L., 2015. Excess body weight increases the burden of age-associated chronic diseases and their associated health care expenditures. Aging (Albany NY) 7 (10), 882.
Belotti, F., Kopinska, J., Palma, A., Piano Mortari, A., 2022. Health status and the great recession. Evidence from electronic health records. Health Econ. 31 (8), 1770–1799.
Bhaskaran, K., dos-Santos-Silva, I., Leon, D.A., Douglas, I.J., Smeeth, L., 2018. Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3·6 million adults in the UK. Lancet Diabetes Endocrinol. 6 (12), 944–953.
Bianchini, E., Brignoli, O., Cricelli, C., Cricelli, I., Lapi, F., Medea, G., Pasqua, A., Pecchioli, S., Piccinni, C., Simonetti, M., Trifirò, G., 2014. VIII Report Health Search. Technical Report, Società Italiana di Medicina Generale, SIMG, available at https://healthsearch.it/documenti/Archivio/Report/VIIIReport_2013-2014/VIII%20Report%20HS.pdf.
Bruzzi, S., Ivaldi, E., Santagata, M., 2022. Measuring regional performance in the Italian NHS: Are disparities decreasing? Soc. Indic. Res. 159 (3), 1057–1084.
Buja, A., Solinas, G., Visca, M., Federico, B., Gini, R., Baldo, V., Francesconi, P., Sartor, G., Bellentani, M., Damiani, G., 2017. Prevalence of heart failure and adherence to process indicators: Which socio-demographic determinants are involved? Int. J. Environ. Res. Public Health 13.
Carpenter, J., Kenward, M., 2012. Multiple Imputation and Its Application. John Wiley & Sons.
Cawley, J., 2015. An economy of scales: A selective review of obesity's economic causes, consequences, and solutions. J. Health Econ. 43, 244–268.
Cawley, J., Meyerhoefer, C., 2012. The medical care costs of obesity: an instrumental variables approach. J. Health Econ. 31 (1), 219–230.
Cawley, J., Meyerhoefer, C., Biener, A., Hammer, M., Wintfeld, N., 2015. Savings in medical expenditures associated with reductions in body mass index among US adults with obesity, by diabetes status. Pharmacoeconomics 33 (7), 707–722.
Colao, A., Lucchese, M., D'Adamo, M., Savastano, S., Facchiano, E., Veronesi, C., Blini, V., Esposti, L.D., Sbraccia, P., 2017. Healthcare usage and economic impact of non-treated obesity in Italy: findings from a retrospective administrative and clinical database analysis. BMJ Open 7, 1–8.
Cricelli, C., Mazzaglia, G., Samani, F., Marchi, M., Sabatini, A., Nardi, R., Ventriglia, G., Caputi, A., 2003. Prevalence estimates for Chronic diseases in Italy: exploring the differences between self-report and primary care databases. J. Public Health Med. 25 (3), 254–257.

Daviglus, M.L., Liu, K., Yan, L.L., Pirzada, A., Manheim, L., Manning, W., Garside, D.B., Wang, R., Dyer, A.R., Greenland, P., et al., 2004. Relation of body mass index in young adulthood and middle age to medicare expenditures in older age. JAMA 292 (22), 2743–2749.

De Silva, A.P., De Livera, A.M., Lee, K.J., Moreno-Betancur, M., Simpson, J.A., 2021. Multiple imputation methods for handling missing values in longitudinal studies with sampling weights: comparison of methods implemented in stata. Biom. J. 63 (2), 354–371.

DeCaria, J.E., Sharp, C., Petrella, R.J., 2012. Scoping review report: obesity in older adults. Int. J. Obes. 36 (9), 1141–1150.

Degli Esposti, E., Sturani, A., Valpiani, G., Di Martino, M., Ziccardi, F., Cassani, A.R., Baraccani, C., Gentile, M., Puglia, M.G., Degli Esposti, L., 2006. The relationship between body weight and drug costs: an Italian population-based study. Clin. Ther. 28 (9), 1472–1481.

Egger, G., Dixon, J., 2014. Beyond obesity and lifestyle: a review of 21st century chronic disease determinants. BioMed Res. Int. 2014, 1–12.

Ezzati, M., Lopez, A.D., Rodgers, A., Vander Hoorn, S., Murray, C.J.L., 2002. Selected major risk factors and global and regional burden of disease. Lancet 360 (9343), 1347–1360.

Fallah-Fini, S., Adam, A., Cheskin, L.J., Bartsch, S.M., Lee, B.Y., 2017. The additional costs and health effects of a patient having overweight or obesity: a computational model. Obesity 25 (10), 1809–1815.

Filippi, A., Vanuzzo, D., Bignamini, A., Sessa, E., Brignoli, O., Mazzaglia, G., 2005. Computerized general practice databases provide quick and cost-effective information on the prevalence of angina pectoris. Italian Heart J. 6 (1), 49–51.

Finkelstein, E.A., Strombotne, K.L., Popkin, B.M., 2010. The costs of obesity and implications for policymakers. Choices 25 (3), 1–6.

Finkelstein, E.A., Trogdon, J.G., Brown, D.S., Allaire, B.T., Dellea, P.S., Kamal-Bahl, S.J., 2008. The lifetime medical cost burden of overweight and obesity: implications for obesity prevention. Obesity 16 (8), 1843–1848.

Flegal, K.M., Graubard, B.I., Williamson, D.F., Gail, M.H., 2005. Excess deaths associated with underweight, overweight, and obesity. JAMA 293 (15), 1861–1867.

Fontaine, K.R., Redden, D.T., Wang, C., Westfall, A.O., Allison, D.B., 2003. Years of life lost due to obesity. JAMA 289 (2), 187–193.

Guh, D.P., Zhang, W., Bansback, N., Amarsi, Z., Birmingham, C.L., Anis, A.H., 2009. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC Public Health 9, 1–20.

Haslam, D.W., James, W.P.T., 2005. Obesity. Lancet 366 (9492), 1197–1209.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–161.

IASO, 2006. Global Obesity Prevalence in Adults. Technical Report, International Association for the Study of Obesity, available at https://http://www.iaso.org.

ISTAT, 2012. Il Diabete in Italia (2012). Technical Report, AIFA - The Medicines Utilisation Monitoring Centre, available at https://www.istat.it/it/files//2012/09/Il-diabete-in-Italia.pdf.

Istat, 2017. Diabetes in Italy. Istituto Nazionale di Statistica. Data retrieved from: https://www.istat.it/en/archivio/202712.

Istat, 2023a. Demographic indicators. Istituto Nazionale di Statistica. Data retrieved from: http://dati.istat.it/Index.aspx?DataSetCode=DCIS_INDDEMOG1&Lang=en.

Istat, 2023b. Estimated resident population - Years 2001–2019. Istituto Nazionale di Statistica. Data retrieved from: http://dati.istat.it/Index.aspx?QueryId=18460&lang=en#.

Kortt, M.A., Langley, P.C., Cox, E.R., 1998. A review of cost-of-illness studies on obesity. Clin. Ther. 20 (4), 772–779.

Lakdawalla, D.N., Goldman, D.P., Shang, B., 2005. The health and cost consequences of obesity among the future elderly. Health Aff. 24 (Suppl2), W5–R30.

Lehnert, T., Heider, D., Leicht, H., Heinrich, S., Corrieri, S., Luppa, M., Riedel-Heller, S., König, H.-H., 2011. Review: health care utilization and costs of elderly persons with multiple chronic conditions. Med. Care Res. Rev. 68 (4), 387–420.

Mazzaglia, G., Ambrosioni, E., Alacqua, M., Filippi, A., Sessa, E., Immordino, V., Borghi, C., Brignoli, O., Caputi, A.P., Cricelli, C., Mantovani, L.G., 2009. Adherence to antihypertensive medications and cardiovascular morbidity among newly diagnosed hypertensive patients. Circulation 120 (16), 1598–1605.

Müller-Riemenschneider, F., Reinhold, T., Berghöfer, A., Willich, S.N., 2008. Health-economic burden of obesity in Europe. Eur. J. Epidemiol. 23 (8), 499–509.

OECD, 2017. OECD Obesity Update 2017. Technical Report, OECD, Paris, France.

OSMED, 2021. National Report on Medicines use in Italy - Year 2021. Technical Report, AIFA, available at https://www.aifa.gov.it/documents/20142/1740782/Rapporto-OsMed-2021.pdf.

OSMED, 2022. National Report on Antibiotics Use in Italy - Year 2022. Technical Report, AIFA - The Medicines Utilisation Monitoring Centre, available at https://www.aifa.gov.it/documents/20142/1664282/Rapporto_Antibiotici_2020.pdf.

Popkin, B.M., Kim, S., Rusev, E.R., Du, S., Zizza, C., 2006. Measuring the full economic costs of diet, physical activity and obesity-related chronic diseases. Obes. Rev. 7 (3), 271–293.

Raleigh, V.S., 2019. Trends in Life Expectancy in EU and Other OECD Countries: Why are Improvements Slowing? Technical Report 108, OECD Health Working Papers, OECD Publishing, Paris.

Reiter, J.P., Raghunathan, T.E., Kinney, S.K., 2006. The importance of modeling the sampling design in multiple imputation for missing data. Surv. Methodol. 32 (2), 143.

Schell, R.C., Just, D.R., Levitsky, D.A., 2021. Methodological challenges in estimating the lifetime medical care cost externality of obesity. J. Benefit-Cost Anal. 12 (3), 441–465.

Sedgwick, P., 2015. Bias in observational study designs: cross sectional studies. BMJ 350, 1–2.

Stevens, J., Cai, J., Thun, M.J., Williamson, D.F., Wood, J.L., 1999. Consequences of the use of different measures of effect to determine the impact of age on the association between obesity and mortality. Am. J. Epidemiol. 150 (4), 399–407.

Thompson, D., Edelsberg, J., Colditz, G.A., Bird, A.P., Oster, G., 1999. Lifetime health and economic consequences of obesity. Arch. Intern. Med. 159 (18), 2177–2183.

Thompson, D., Edelsberg, J., Kinsey, K.L., Oster, G., 1998. Estimated economic costs of obesity to U.S. business. Am. J. Health Promot. 13 (2), 120–127.

Tsai, A.G., Williamson, D.F., Glick, H.A., 2011. Direct medical cost of overweight and obesity in the USA: a quantitative systematic review. Obes. Rev. 12 (1), 50–61.

Tucker, D.M., Palmer, A.J., Valentine, W.J., Roze, S., Ray, J.A., 2006. Counting the costs of overweight and obesity: modeling clinical and cost outcomes. Curr. Med. Res. Opin. 22 (3), 575–586.

von Lengerke, T., Krauth, C., 2011. Economic costs of adult obesity: a review of recent European studies with a focus on subgroup-specific costs. Maturitas 69 (3), 220–229.

von Lengerke, T., Reitmeir, P., John, J., 2006. Direct medical costs of (severe) obesity: a bottom-up assessment of over- vs. normal-weight adults in the KORA-study region (Augsburg, Germany). Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany)) 68 2, 110–115.

Withrow, D., Alter, D.A., 2011. The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. Obes. Rev. 12 (2), 131–141.

Wolf, A.M., Colditz, G.A., 1998. Current estimates of the economic cost of obesity in the United States. Obes. Res. 6 (2), 97–106.

Yang, Z., Hall, A.G., 2008. The financial burden of overweight and obesity among elderly Americans: the dynamics of weight, longevity, and health care cost. Health Serv. Res. 43 (3), 849–868.