Breast Cancer Res Treat
DOI 10.1007/s10549-017-4210-z

EPIDEMIOLOGY

Adherence to guidelines and breast cancer patients survival: a population-based cohort study analyzed with a causal inference approach

Anita Andreano[1,2] · Paola Rebora[1] · Maria Grazia Valsecchi[1] · Antonio Giampiero Russo[2]

Received: 12 December 2016 / Accepted: 17 March 2017

Abstract
Purpose There is a lack of real-world studies evaluating the impact on survival of an evidence-based pathway of care in breast cancer. The aim of this work is to investigate the effect of adherence to guidelines on long-term survival for a cohort of Italian breast cancer patients.
Methods The cohort included incident female breast cancer cases (2007–12**),** from the registry of the Milan province (Italy), not metastatic at diagnosis and receiving primary surgery. We selected sets of indicators, according to patient and tumor characteristics. We then defined the pathway of care as adherent to guidelines if it fulfilled at least 80% of the indicators. Indicators were measured using different administrative health databases linked on a unique key. A causal inference approach was used, drawing a directed acyclic graph and fitting an inverse probability weighted marginal structural model, accounting for patient's demographic, socioeconomic and tumor characteristics.
Results The analysis included 6333 patients, 69% of them were classified as having an adherent care. Mean age was 61 years (standard deviation, 13.6 years) and half of the patients were in Stage I (50%) at diagnosis. Median follow-up time was 5.6 years. Overall, 5-year survival was 90% (95% CI, 89–91%). The estimated risk of death was 30% lower for patients with adherent than nonadherent care (hazard ratio [HR], 0.66; 95% CI, 0.55–0.77).
Conclusions Our study confirms, in real-world care, the impact on survival of receiving a care pathway adherent to guidelines in non-metastatic breast cancer patients.

Keywords Breast cancer care · Adherence to guidelines · Survival · Causal inference · Administrative health databases · Process indicators

Abbreviations

CI	Confidence interval
DAG	Directed acyclic graph
HR	Hazard ratio
IPW	Inverse probability weighted
RCT	Randomized controlled trial

Introduction

Evidence-based guidelines, including those for cancer, provide recommendations to improve the quality of medical care and ultimately patient outcomes [1]. Whenever possible, guidelines should be derived from randomized controlled trials (RCTs) and meta-analyses on RCTs, as they represent the most reliable evidence [2–4]. Nevertheless, RCTs are often conducted in third level centers, usually excluding the elderly and people with serious comorbidities. Thus, after the establishment of guidelines, their impact on survival should be confirmed in clinical practice. Pragmatic clinical trials are an attempt in this direction, but they are generally not designed to evaluate an entire pathway of care [5]. On the contrary, routinely collected healthcare data and

Electronic supplementary material The online version of this article (doi:10.1007/s10549-017-4210-z) contains supplementary material, which is available to authorized users.

✉ Anita Andreano
anita.andreano@unimib.it

1 Center of Biostatistics for Clinical Epidemiology, School of Medicine and Surgery, University of Milan Bicocca, Via Cadore 48, 20900 Monza, Italy

2 Epidemiology Unit, Agency for Health Protection of the Province of Milan, Milan, Italy

cancer registries offer a possibility to perform real-world studies, analyzing pathways of care without the need for large economical investments. There are many studies examining the effect of adherence to guidelines during hospitalization for acute diseases on survival [6, 7]. Conversely, this has rarely been done for complex pathways in chronic diseases, including tumors [8–11]. The motive is the complexity of evaluating simultaneously the impact of multiple procedures performed over an extended time-period. Moreover, these studies are observational by nature and thus prone to confounding [12]. Consequently, it is essential to collect all possible confounders—not only those related to tumor aggressiveness [4, 13], but also those relevant to socioeconomic status- and to apply a sound methodological approach to evaluate the survival benefit of an evidence-based care pathway. We took a causal inference approach, that allows to estimate a causal effect of adherence to guidelines on survival, instead of the mere association, even in an observational study [14, 15].

We investigated the effect of receiving diagnosis and care adherent to guidelines on long-term survival in a cohort of breast cancer patients, identified through population cancer registers. Indicators of adherence were based on administrative health data. The causal analysis accounted for patient demographics, socioeconomic status and tumor characteristics.

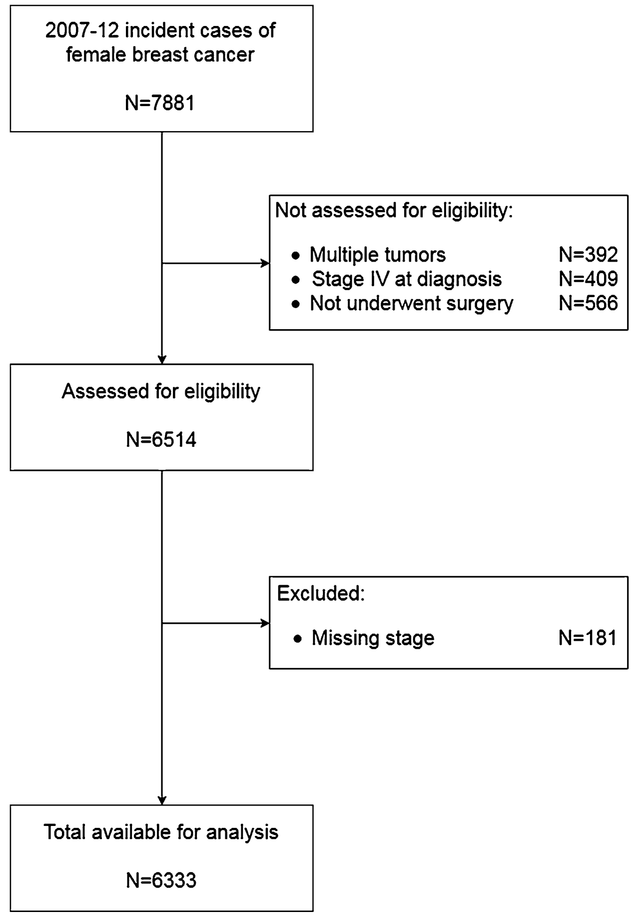


Fig. 1 Flow chart

Methods

Study population and definition of clinical pathways

This is a retrospective cohort study. We assessed for eligibility all women with an epithelial breast cancer diagnosed during 2007–2012 and included in the population cancer registry of the Milan province (Lombardy, Italy). Exclusion criteria were: having had a previous malignant tumor (International Classification of Diseases—9th revision -Clinical Modification codes 140-208), being in stage IV, not having undergone surgery. The first criteria avoids to include patients with administrative data contaminated from procedures due to the treatment of the first primary cancer (non-breast), while the remaining exclusions are motivated by the lack of well-defined guidelines [16] (Fig. 1).

The study population was divided in seven groups, each expected to follow a different pathway, on the basis of having received neoadjuvant treatment or not, stage at diagnosis (I vs. II-III, American Joint Committee on Cancer System), and type of surgery (conservative vs. radical). Four to seven process indicators were selected for each group, based on recommendations from clinical guidelines [17, 18] and on strength of the association between survival and the measured procedure, as reported in the literature.

The seven indicators were: (1) An interval from diagnosis to breast surgery shorter than 3 months, because an increase in time to surgery was shown to enhance the risk of death (hazard ratio of 1.1 every 30 days up to 120 days) [19]. We used as time of diagnosis the mammography date or, for women under 50 years, the breast sonography date. (2) Whether patients had lymph-node assessment, because understaging at diagnosis has a therapeutic impact, since it prevents patients from receiving adequate therapies, e.g. adjuvant therapy [18, 20]. (3) Whether patients underwent cytological and/or histological assessment before treatment, as this is necessary to define grade and molecular subtype [21] and consequently plan an appropriate treatment. (4) Whether patients undergoing breast-conserving surgery received a following radical surgery within 3 months, as a proxy for positive margins [22]. (5) In patients with pathological stage II-III, whether they had received adjuvant therapy (chemotherapy with or without hormone therapy below age 70, and chemo or hormone therapy for patients older than 70 years), because it is established that this ensures a significant reduction in mortality [23–25]. (6) Analogously, whether patients receiving breast-conserving surgery had a breast radiation treatment within a year after surgery [26], or if patients with a primary tumor extent $T \geq 3$ and undergoing neoadjuvant treatment received radiotherapy regardless of the type of surgery [27]. (7) Whether patients had mammography follow-up within 18 months after surgery [28, 29]. The indicators considered for each group are detailed in Table 1.

Definition of adherence and potential confounders

Firstly, we counted the number of indicators that each patient had met and divided it by the total number for which the patient was eligible, obtaining the proportion of adherence. Follow-up indicators were not counted for patients who had died or developed metastasis in the time window used to calculate them. Metastasis onset was assessed by administrative data. For the primary analysis, we defined as adherent the care pathway of patients with a proportion of met indicators equal or greater than 80%. Sensitivity analyses were planned at different cut-offs, and considering adherence as an ordinal and a continuous variable (Appendix 1).

We constructed a directed acyclic graph (DAG) [30, 31] to represent assumptions regarding the underlying causal relationships between guideline adherence, survival and a set of clinical and socioeconomic variables. The DAG (Online Resource 1) utilizes these assumptions to select the potential confounders, rather than relying on the statistical associations observed in the data at hand. The selected confounders are then used in the statistical analysis aiming at evaluating the ‘causal’ impact of receiving a care adherent to guidelines on survival.

Primary outcome measure

The primary outcome was overall survival. Patients moving outside the Lombardy region were lost to follow-up and censored at the last available contact. Administrative censoring was set at 31/06/2015. We first investigated the relationship between adherence and survival in the entire cohort. We then performed a subgroup analysis for age (<50 vs. 50–69 vs. >69). We also explored the association between adherence and survival within clinical pathways,

Table 1 Percentage of patients who fulfilled each included indicator (row) by the assignment criteria (columns) of the 7 clinical pathways, Milan province, Italy, 2007–12

		No neoadjuvant therapy				Neoadjuvant therapy			All
		Conservative surgery		Radical surgery		T stage < 3		T stage ≥ 3	
		Stage I	Stage II–III	Stage I	Stage II–III	Conservative surgery	Radical surgery		
Indicators		N = 2586	N = 1780	N = 322	N = 1117	N = 307	N = 168	N = 53	N = 6333
1	Mammography (or ultrasound < 50 years) within 3 months before surgery	81.28	83.31	85.71	79.59	79.48	70.24	41.51	81.07
2	Lymph node staging within 3 months before surgery[a]	88.94	89.21	81.37	70.64	90.55	63.10	71.70	84.65
3	Cyto-histologic assessment within 3 months before surgery[b]	70.73	69.38	70.50	71.80	75.57	76.19	73.58	70.93
4	No second surgery within 3 months	97.52	94.94	n.i.	n.i.	97.06	n.i.	n.i.	96.51
5	Adjuvant medical treatment[c]	n.i.	75.17	n.i.	77.72	n.i.	n.i.	n.i.	76.14
6	Radiotherapy after conservative surgery[d]	84.90	84.38	n.i.	n.i.	81.25	n.i.	68.00	84.29
7	Follow-up mammography in the 18 months after surgery	88.83	85.90	80.94	71.80	88.93	76.28	71.11	84.27
	N of indicators	6	7	4	5	6	4	5	

n.i. not included

[a] Or at surgery. Lymph node staging within 3 months before the start of neoadjuvant therapy and surgery for patients undergoing neoadjuvant chemotherapy

[b] Or neoadjuvant treatment, for patients undergoing it

[c] Chemotherapy/hormone therapy for patients > 70 years; chemotherapy ≤ 70 years

[d] Or any surgery after neoadjuvant therapy in patients with T stage ≥ 3

i.e., groups of patients eligible to the same procedures and thus with the same indicators. We did it for the three most frequent groups: patients not undergoing neo-adjuvant treatment and undergoing conservative surgery in stage I (pathway 1) and stage II–III (pathway 2), and undergoing radical surgery in stage II–III (pathway 4).

Sources of data

The data used to calculate the process indicators came from electronic sources of health data, including hospital discharge database, prescription database, database of outpatient diagnostic, and therapeutic procedures. Deterministic record linkage on a unique key was used to match all information at patient level (further details can be found in a previous paper [32]). Record linkage was performed at the local health authority, which owns the health data and houses the cancer registry. Patient characteristics and data on breast cancer at diagnosis were available from the cancer registry of the Milan province, and included age, pathological stage, grading, molecular classification [21], and number of positive lymph nodes. We also considered indices of socioeconomic status derived from administrative health databases (marital status, employment, education), calculated the Charlson comorbidity index from inpatient and outpatient databases [33, 34], and included deprivation index on a zip code basis [35].

We derived the patient vital status from the health registry office of the entire Lombardy region [36], where an update is performed every 6 months covering at least 95% of deaths.

Statistical analysis

Differences between the distribution of the covariates among patients with adherent ($\geq$80%) and nonadherent care were assessed using the χ^2 test. To investigate the 'causal' relationship between survival and adherence (dichotomous variable with a 80% cut-off), we fitted an inverse probability weighted (IPW) marginal structural model [37]. After selecting potential confounders based on the DAG, we created a pseudo-population by the use of (stabilized) inverse probability of adherence weights to mitigate the differences between patients with adherent and nonadherent care. These weights were estimated through a multivariable logistic regression model on adherence, including as covariates the potential confounders: molecular type, stage, grade, their first-degree interactions, age (always as a continuous variable with restricted cubic spline functions), Charlson comorbidity index, number of positive nodes, marital status, employment, education, and deprivation index. The final weights were obtained by multiplying these weights with inverse probability of censoring weights, obtained by an analogous logistic model on censoring. The distribution of stabilized weights is fairly symmetrical, it ranges from 0.328 to 7.184 and is presented in Appendix 2. In order to evaluate the balance induced by these weights, the confounders among patients with adherent and nonadherent care in the pseudo-population were compared by standardized differences of the mean/proportion. We fitted an IPW logistic model on a person-time dataset, which estimates an odds ratio approximating a discrete-time hazard ratio (HR) [37]. This model is analogous to an IPW Cox model but it allows more flexibility. We fitted a model assuming a constant hazard ratio and a second one allowing for a time-varying HR including an interaction term between adherence and a flexible function of time. After exploring different time functions, we chose a restricted cubic spline with three knots at 18, 36, and 68 months (corresponding to the 30th, 60th, and 90th percentiles) on Akaike Information Criteria basis.

Sensitivity analyses were performed considering different cut-offs of adherence (from 60 to 90% by 10%, Appendix 1.1), considering adherence as a four classes ordinal variable and as a continuous variable (Appendix 1.2).

Percentages of missing values were 13.6% for grading, 13.7% for molecular subtypes, 7.6% for number of positive nodes, 16.5% for marital status, 24.3% for education, and 0.6% for employment. All analyses were performed after multiple imputations on missing values, assuming they were missing at random given age, stage, type of surgery, and deprivation index. Imputation was performed by fitting a model on each variable with missing values including adherence, the indicators common to all pathways, and all potential confounders (50 imputed datasets, multivariate imputation by fully conditional specification [38], MI SAS procedure). Parameter estimates were combined according to Rubin using the MIANALYZE [39] procedure of the software SAS. As sensitivity analyses, we estimated the proportional hazards IPW marginal structural model also on the 3238 patients with all necessary data on potential confounders in the original dataset (Appendix 1.3) and on patients $\leq$69 years (Appendix 1.4).

Results are presented both as HRs with their 95% confidence intervals (CI), and as survival curves. Tests were two-sided. All analyses were performed using SAS software v.9.4 [40]. Graphs were performed using R [41].

Results

Study population characteristics

Between 2007 and 2012, 7881 incident cases of female epithelial breast cancer were recorded in the cancer registry. Based on pre-defined exclusion criteria, not included patients were: 392 because of multiple tumors, 409 for being in stage IV, and 566 for not receiving surgery. We also excluded 181 patients with missing data on stage, as this variable is essential to assign the clinical pathway (Fig. 1). All remaining patients (6333) were included in the analysis. Sixty-nine percent (4374) of the population was classified as having an adherent care at the 80% cut-off.

Median follow-up time was 5.6 years (95% CI, 5.4–5.7 years). Seventy-one patients (1.1%) emigrated outside Lombardy during the study period and were censored (median follow-up time, 3 years). Overall, 5-year survival was 90% (95% CI, 89–91%).

Among the seven groups, the one including women in stage I undergoing conservative surgery was the most frequent (2586, Table 1), while the pathway starting with neoadjuvant therapy in patients with pT $\geq$ 3 had only 53 patients. Among indicators common to all groups, cyto-histologic assessment before surgery (indicator 3) was the more stable varying from 69.4 to 76.2% (SD, 2.5).

Demographic and socioeconomic characteristics of the population and features of the tumor at diagnosis are described in Table 2. Mean age at diagnosis was 61 years (SD, 13.6), and 46.3% of patients were in the screening age range (50–69 years). There were more patients aged 50–69 years and less patients aged 70 years and older in the adherent group with respect to nonadherent (48.9 vs. 40.6% and 25.3 vs. 34.0%). Overall, grade 2 was the most represented at diagnosis (53.5%), and there were less grade 3 patients among adherents (33.8 vs. 37.1%). The majority of patients were in Stage I (49.7%), and there were less advanced stages in the adherent group (Stage II–III, 47.4 vs. 56.8% in nonadherent). Concerning molecular subtype, luminal A and B were the most frequent accounting for 37.7 and 38.7%. At least one comorbidity was found for 21.7% of patients. Patients with 3 or more comorbidities were 3.4% in the nonadherent vs. 1.4% in the adherent group. Thirty-eight percent of patients lived in a town-hall in the richest quintile of the Lombardy region (deprivation index V), with no differences between adherent and non-adherent patients. Almost seventy percent of patients were married, 72.8% among adherent and 63.1% among non-adherent. Overall, 60.1% had a middle school degree or lower, with the greatest difference between adherent and nonadherent for middle school (26.1 vs. 23.2%). Concerning occupational status, 40.0% were retired; among the remaining, 51.7% (1974/3817) were housewife or unemployed, 52.0% among adherents and 50.9% among non-adherents. After imputation of variables with missing values, the distribution of patients in classes and adherence groups remained similar and associations between potential confounders and adherence groups were confirmed (Table 2). After weighting, the standardized differences of the mean/proportion in confounders between adherent and nonadherent were all lower than 0.025 (Fig. 2, results shown for the first imputation).

Association between adherence and survival

Overall, the estimated risk of death was 30% lower for patients receiving adherent care compared to those with nonadherent care (HR, 0.66; 95% CI, 0.55–0.77, Table 3). The model not assuming a constant HR over time, showed a decreasing protective effect over the first 5 years of follow-up (Table 3 and Fig. 3, panel a). The model-based estimate of 5-year survival was 93.2% for adherent vs. 83.4% for nonadherent patients (Fig. 3, panel b). Regarding the HRs of death for adherent vs. nonadherent patients, in the subgroup analyses stratifying for age (Table 3), the estimated HRs were 0.88 (95% CI, 0.55–1.38) < 50 years, 0.68 (95% CI, 0.49–0.95) between 50–69 years and 0.61 (95% CI, 0.48–0.77) > 69 years. The trend of the HR over time for the three age subgroups is shown in Table 3 and Online Resource 2.

Table 4 shows results in the three most frequent pathways, all characterized by no neoadjuvant therapy. For patients in pathway 1 (stage I and conservative surgery), the reduction of death risk for receiving adherent care was larger than 50% (HR, 0.44; 95% CI, 0.28–0.69, Table 4). In pathway 2 (stages II–III and conservative surgery), the HR was 0.66 (95% CI, 0.49–0.90). For pathway 4 (stages II–III and radical surgery), the HR was 0.70 (95% CI, 0.53–0.92).

Results were consistent across the performed sensitivity analyses (Appendix 1).

Discussion

In this population-based study of women diagnosed with non-metastatic breast cancer, patients whose care was adherent to a pre-defined set of guidelines were 30% less likely to die in the first 5 years after diagnosis compared to nonadherent ones.

Two previous studies investigated the relationship between adherence and survival for breast cancer using a conditional multivariable Cox model, accounting for age and tumor characteristics, but not for socioeconomic factors and comorbidities as potential confounders [11, 13]. The study of Cheng et al. [13] found a better overall survival for patients with a 100% adherent care (HR of death for adherent vs. nonadherent, 0.46). The study from van de Water [11] found analogous results, both below 65-year (HR, 0.57) and over 75-year of age (HR, 0.62).

Few other real-world studies have attempted to relate adherence to guidelines with survival in tumors [8–10]. However, due to controversial aspects of this type of studies, heterogeneous analytical methodologies were used. One issue is the rational used to choose indicators to be included in each pathway. The aim of the present study was

Table 2 Demographic and socio-economic characteristics of the population and features of the tumor at diagnosis, Milan province, Italy, 2007–12

		All cases				Imputed dataset (I = 50)		
		Overall N = 6333 No. (%)	Adherent N = 4374 No. (%)	Nonadherent N = 1959 No. (%)	χ^2 test *P*	Adherent N = 4374 No. (%)	Nonadherent N = 1959 No. (%)	χ^2 test *P*
Age								
	No.miss(%)	0			<.001			
	<50 years	1626 (25.7)	1129 (25.8)	497 (25.4)		n.i.	n.i.	
	50–69 years	2933 (46.3)	2138 (48.9)	795 (40.6)		n.i.	n.i.	
	>69 years	1774 (28.0)	1107 (25.3)	667 (34.0)		n.i.	n.i.	
Grading								
	No.miss(%)	859 (13.6)	619 (14.1)	240 (12.3)	<.001			0.023
	1	636 (11.6)	447 (11.9)	189 (11.0)		509 (11.6)	218 (11.1)	
	2	2929 (53.5)	2037 (54.3)	892 (51.9)		2397 (54.8)	1014 (51.8)	
	3	1909 (38.9)	1271 (33.8)	638 (37.1)		1468 (33.6)	727 (37.1)	
Stage								
	No.miss(%)	0			<.001			
	I	3145 (49.7)	2299 (52.6)	846 (43.2)		n.i.	n.i.	
	II	2287 (36.1)	1515 (34.6)	772 (39.4)		n.i.	n.i.	
	III	901 (14.2)	560 (12.8)	341 (17.4)		n.i.	n.i.	
Molecular subtype								
	No.miss(%)	867 (13.7)	620 (14.2)	247 (12.6)	0.780			0.736
	Luminal A	2063 (37.7)	1418 (37.8)	645 (37.7)		1660 (38.0)	739 (37.7)	
	Luminal B	2114 (38.7)	1455 (38.8)	659 (38.5)		1702 (38.9)	752 (38.4)	
	Luminal Her2	554 (10.1)	385 (10.2)	169 (9.9)		437 (10.0)	192 (9.8)	
	Her2 type	291 (5.3)	190 (5.1)	101 (5.9)		224 (5.1)	117 (6.0)	
	Triple negative	444 (8.1)	306 (8.1)	138 (8.0)		351 (8.0)	159 (8.1)	
N of positive nodes								
	No.miss(%)	483 (7.6)	372 (8.5)	111 (6.7)	0.049			0.035
	0	3772 (64.5)	2624 (65.6)	1148 (62.1)		2798 (64.0)	1194 (61.0)	
	1–3	1374 (23.5)	921 (23.0)	453 (24.5)		1020 (23.3)	471 (24.0)	
	4–9	392 (6.7)	258 (6.4)	134 (7.3)		330 (7.5)	164 (8.4)	
	≥10	312 (5.3)	199 (5.0)	113 (6.1)		226 (5.2)	130 (6.6)	
Charlson index								
	No.miss(%)	0			<.001			
	0	4957 (78.3)	3471 (79.4)	1486 (75.9)		n.i.	n.i.	
	1–2	1246 (19.7)	840 (19.2)	406 (20.7)		n.i.	n.i.	
	≥3	130 (2.0)	63 (1.4)	67 (3.4)		n.i.	n.i.	
Deprivation index								
	No.miss(%)	0			0.556			
	I	864 (13.6)	576 (13.2)	288 (14.7)		n.i.	n.i.	
	II	770 (12.2)	529 (12.1)	241 (12.3)		n.i.	n.i.	

Table 2 continued

		All cases				Imputed dataset (I = 50)		
		Overall N = 6333 No. (%)	Adherent N = 4374 No. (%)	Nonadherent N = 1959 No. (%)	χ^2 test *P*	Adherent N = 4374 No. (%)	Nonadherent N = 1959 No. (%)	χ^2 test *P*
	III	872 (13.8)	607 (13.9)	265 (13.5)		n.i.	n.i.	
	IV	1445 (22.8)	1009 (23.0)	436 (22.3)		n.i.	n.i.	
	V	2382 (37.6)	1653 (37.8)	729 (37.2)		n.i.	n.i.	
Marital status								
	No.miss(%)	1043 (16.5)	714 (16.3)	329 (16.8)	<.001			<.001
	Never married	516 (9.8)	343 (9.4)	173 (10.6)		419 (9.6)	203 (10.4)	
	Married	3693 (69.8)	2664 (72.8)	1029 (63.1)		3203 (73.2)	1258 (64.2)	
	Divorced	205 (3.9)	139 (3.8)	66 (4.1)		182 (4.2)	84 (4.3)	
	Widow	876 (16.6)	514 (14.0)	362 (22.2)		570 (13.0)	414 (21.1)	
Education								
	No.miss(%)	1540 (24.3)	1036 (23.7)	504 (25.7)	0.025			<.001
	None or elementary school	1669 (34.8)	1134 (34.0)	535 (36.8)		1758 (40.2)	863 (44.1)	
	Middle School	1209 (25.3)	871 (26.1)	338 (23.2)		1122 (25.6)	424 (21.6)	
	High School	1684 (35.1)	1185 (35.5)	499 (34.3)		1267 (29.0)	541 (27.6)	
	College and higher	231 (4.8)	148 (4.4)	83 (5.7)		227 (5.2)	131 (6.7)	
Employment								
	No.miss(%)	37 (0.6)	25 (0.6)	12 (0.6)	0.007			0.006
	Housewife/unemployed	1974 (31.3)	1398 (32.2)	576 (29.6)		1404 (32.1)	579 (29.6)	
	Workers	488 (7.8)	360 (8.3)	128 (6.6)		361 (8.2)	128 (6.5)	
	Office worker/teachers	1088 (17.3)	753 (17.3)	335 (17.2)		762 (17.4)	338 (17.2)	
	Managers and professionals	230 (3.6)	150 (3.4)	80 (4.1)		152 (3.5)	82 (4.2)	
	Retired	2516 (40.0)	1688(38.8)	828 (42.5)		1695 (38.8)	832 (42.5)	

The average across all imputations is presented for imputed datasets

I number of imputations, *n.i.* not included

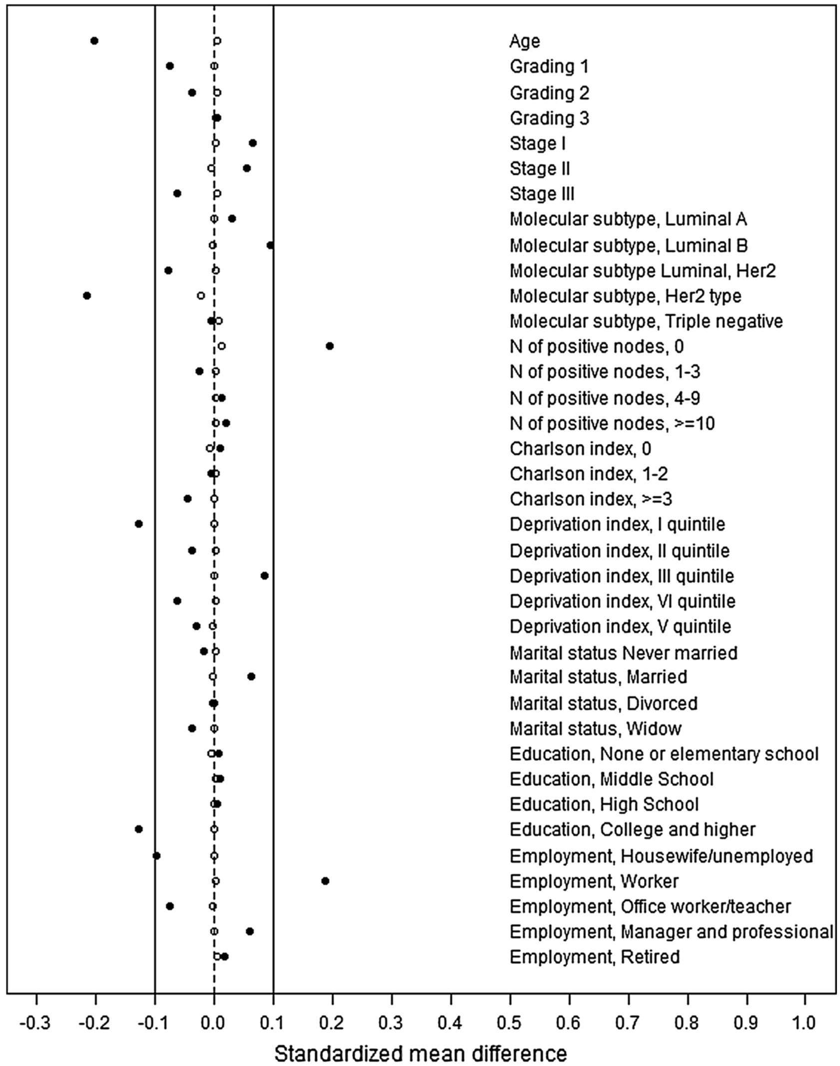


Fig. 2 Standardized differences between adherent and nonadherent patients of the mean/proportion of confounders in the population (*filled circles*) and in the pseudo-population created using adherence weights (*empty circles*). In the pseudo-population, the difference in the distribution of confounders between adherent and nonadherent patients is reduced

Table 3 Results from the inverse probability weighted marginal structural models investigating the relationship between adherence and survival in the whole population and stratifying for age

	No.	% of adherents	Proportional hazard model HR (95% CI)	Nonproportional hazard model		
				HR (time since diagnosis) (95% CI)		
				1 year	2 years	5 years
All patients	6333	69	0.66 (0.55–0.77)	0.47 (0.36–0.61)	0.56 (0.47–0.68)	0.90 (0.70–1.15)
Age						
<50 years	1626	69	0.88 (0.55–1.38)	0.63 (0.27–1.38)	0.61 (0.35–1.05)	1.34 (0.67–2.70)
50–69 years	2933	73	0.68 (0.49–0.95)	0.51 (0.30–0.87)	0.59 (0.40–0.86)	0.91 (0.56–1.49)
>69 years	1774	62	0.61 (0.48–0.77)	0.45 (0.31–0.67)	0.56 (0.42–0.74)	0.77 (0.55–1.08)

HR hazard ratio, *CI* confidence interval

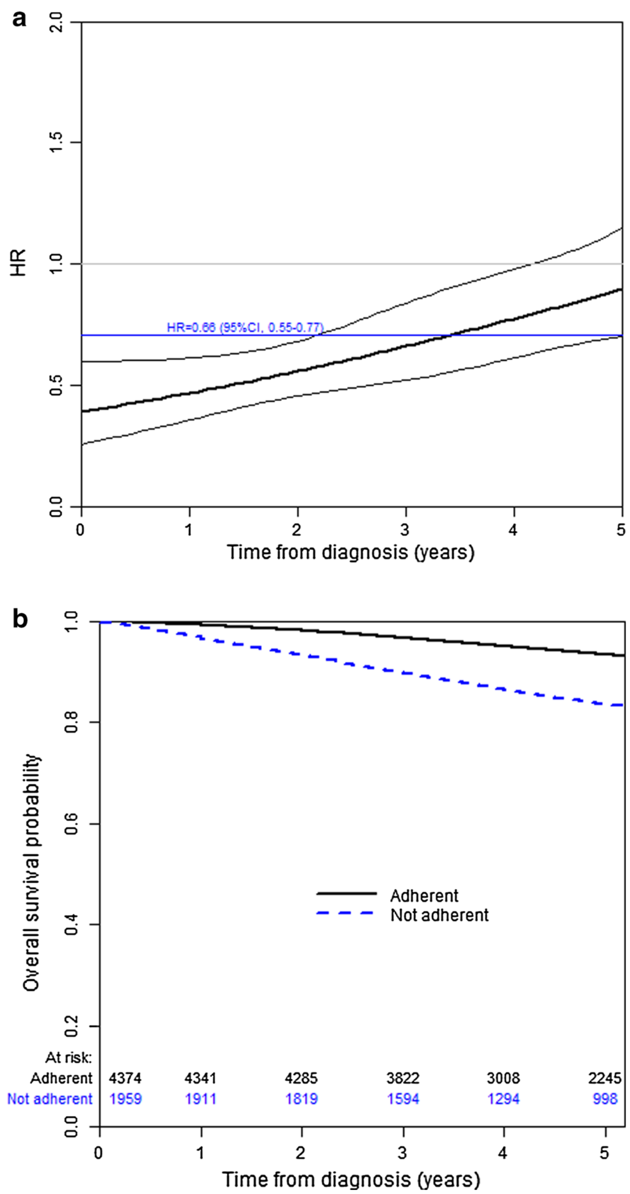


Fig. 3 Results from the inverse probability weighted (IPW) marginal structural model for the association between adherence and survival. Hazard ratio (HR) of death for adherent vs. nonadherent patients over time (*panel a*) and survival curves for the two groups (*panel b*)

to confirm the benefit of adherence to guidelines—which already refer to procedures that demonstrated a survival benefit in RCTs—on survival in real-world care. Consequently, we only included indicators measuring procedures known or expected to increase directly survival. We did not include indicators measuring different aspects of the care pathway, such as quality of life or excess of staging/follow-up procedures [42]. A second controversial is if, and how, to weight the different indicators. Since there is no widely accepted method to objectively assign loads, and subjectively defined weights may reduce generalizability [43, 44], we assigned the same weight to all indicators included in each pathway. The last issue is how to determine which degree of adherence is necessary to see a benefit [45]: we performed sensitivity analyses, both at different cut-offs and considering adherence as an ordinal and continuous variable, and found that results were robust to the definition of adherence.

This is an observational study, consequently prone to confounding by the biology of the cancer and patient characteristics. To account properly for that, we took a causal inference approach, hypothesizing a network of causal relationships in the DAG. All identified potential confounders, both patient-related and tumor-related, were available and accounted for in the analysis. The IPW model weights each patient by (the inverse of) its propensity to adhere/not adhere to guidelines with respect to all confounders. This weighting creates a new population (pseudo-population) in which confounders are balanced between the two groups (Fig. 2).

In the elderly, overall survival could include a substantial number of not breast cancer related deaths, also in the first five years after diagnosis. However, both adherent and nonadherent patients are subjected to mortality from other causes with no expected major differences among them, especially after adjusting for the major determinants of non-related breast cancer death in the elderly (i.e., age, comorbidities and socio-economic status) [46]. Moreover, the results of the sensitivity analysis on patients younger

Table 4 Results from the inverse probability weighted marginal structural models investigating the relationship between adherence and survival in the three most frequent paths (all no neoadjuvant therapy)

Pathway		No.	% of adherents	HR (95% CI)	*P*
1	Conservative surgery, stage I	2586	77	0.44 (0.28–0.69)	<.001
2	Conservative surgery, stage II–III	1780	69	0.66 (0.49–0.90)	0.008
4	Radical surgery, stage II–III	1117	63	0.70 (0.53–0.92)	0.011

HR hazard ratio, *CI* confidence interval than 70 years of age (Appendix 1.4) are consistent with those of the entire sample.

This study confirms, at a population level, the impact of adherence to guidelines on survival in non-metastatic breast cancer patients undergoing surgery. This finding should further increase the awareness of the healthcare professionals taking care of women with breast cancer on the importance of adhering to evidence-based guidelines.

Funding Supported by grant RF-2011-02348959 from the Italian Ministry of Health.

Compliance with ethical standards

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

Ethical standards All analyses were performed in compliance with the Italian law on health data protection.

Appendix 1: Sensitivity analyses

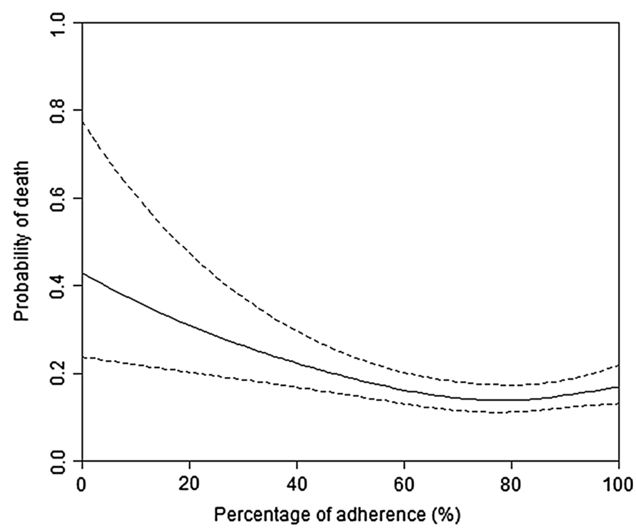
1. Results of the IPW marginal structural model, assuming proportional hazards, investigating the association between survival and adherence as a dichotomous variable, using different cut-offs to define adherent care. Results, here and in the following section, are presented as the Hazard ratio (HR) of death for adherent vs. nonadherent patients and their 95% confidence intervals (CI).

Cut-off	% of adherent pts	HR (95% CI)	*P*
60%	84	0.71 (0.58–0.87)	0.001
70%	77	0.63 (0.52–0.76)	<.001
80% (main analysis)	69	0.66 (0.55–0.77)	<.001
90%	38	0.88 (0.73–1.07)	0.21

2. Results of the IPW marginal structural model, assuming proportional hazards, investigating the association between survival and adherence as a four class ordinal variable (cut-points: 20th percentile, corresponding to 67%; 30th percentile, 75%; 60th percentile, 86%), and as a continuous variable. For the ordinal variable, probability of adherence given potential confounders $f(A|\mathbf{L})$, was estimated using a cumulative logistic regression model with the same covariates of the main analysis, with the exception of all first-degree interactions.

Adherence	HR (95% CI)	*P*
≤67% vs. > 86%	2.83 (2.29–3.45)	<.001
>67 and ≤ 75% vs. > 86%	1.57 (1.14–2.14)	0.005
>75 and ≤ 86% vs. > 86%	1.01 (0.80–1.28)	0.911

For adherence treated as a continuous variable varying between 0 and 100, $f(A|\mathbf{L})$ is a probability density function and it was estimated through a linear regression model, including the same covariates of the main analysis, and then assuming a normal distribution with constant variance. We estimated an IPW marginal structural model for the association between survival and adherence as a flexible function (restricted cubic spline with three knots at 50, 85 and 90%). The graph illustrates the 8-year probability of death for increasing adherence level, which is higher for low adherence percentage and decreases almost linearly up to 80%. This trend motivated the choice of the cut-off for the main analysis presented in the article.



3. Complete cases analysis: we estimated an IPW marginal structural model for the association between survival and adherence assuming proportional hazards using only patients with all necessary potential confounders in the original dataset (3238). Characteristics of this population are described in the table below. The proportion of adherent patients was 69%. The estimated HR was 0.71 (95%CI, 0.57-0.90, $P = 0.004$).

	No. of pts 3238			
	Adherent No. 2237		Nonadherent No. 1001	
	N	%	N	%
Age				
<50 years	654	29.24	298	29.77
50–69 years	1037	46.36	362	36.16
>69 years	546	24.41	341	34.07
Grading				
1	241	10.77	99	9.89
2	1189	53.15	523	52.25
3	807	36.08	379	37.86
Stage				
1	1143	51.10	439	43.86
2	780	34.87	402	40.16
3	314	14.04	160	15.98
Molecular subtype				
Luminal A	817	36.52	375	37.46
Luminal B	897	40.10	377	37.66
Luminal Her2	240	10.73	117	11.69
Her2 type	96	4.29	56	5.59
Triple negative	187	8.36	76	7.59
N of positive nodes				
0	1431	63.97	618	61.74
1–3	524	23.42	260	25.97
4–9	161	7.20	66	6.59
≥10	121	5.41	57	5.69
Charlson index				
0	1733	77.47	755	75.42
1–2	471	21.05	206	20.58
≥3	33	1.48	40	4.00
Deprivation index, quintiles				
I	273	12.20	146	14.59
II	265	11.85	128	12.79
III	336	15.02	135	13.49
IV	525	23.47	208	20.78
V	838	37.46	384	38.36
Marital status				
Never married	238	10.64	113	11.29
Married	1627	72.73	635	63.44
Divorced	88	3.93	47	4.70
Widow	284	12.70	206	20.58
Education				
None or elementary school	713	31.87	370	36.96
Middle School	606	27.09	230	22.98
High School	814	36.39	339	33.87
College and higher	104	4.65	62	6.19
Employment				
Housewife/unemployed	553	24.72	211	21.08
Workers	242	10.82	90	8.99
Office worker/teachers	540	24.14	232	23.18
Managers and professionals	86	3.84	48	4.80
Retired	816	36.48	420	41.96

4. Results of the IPW marginal structural model, assuming proportional and non-proportional hazards, investigating the association between survival and adherence as a dichotomous variable, excluding patients > 69 years.

No.	% of adherents	Proportional hazard model HR (95% CI)	Non proportional hazard model HR(time since diagnosis) (95% CI)		
			1 year	2 years	5 years
4559	72	0.73 (0.56–0.95)	0.55 (0.36–0.86)	0.58 (0.43–0.80)	1.04 (0.70–1.56)

HR hazard ratio, *CI* confidence interval

Appendix 2. Distribution of weights

Distribution of stabilized weights derived from the analysis of 50 imputed datasets (missing data imputed by fully conditional specification using MI SAS procedure). Standard errors (SE) have been pooled according to Rubin [39]. Minimum and maximum are the lowest and highest values found across all datasets.

Stabilized weights Inverse probability of:	Mean	SE	Median	I Q	III Q	Min	Max
Adherence	1.005	0.001	0.958	0.888	1.079	0.328	7.177
Censoring	1.000	0.000	1.000	0.999	1.000	0.949	1.123
Adherence and censoring	1.005	0.001	0.958	0.888	1.079	0.328	7.184

Q quartile, *Min* minimum, *Max* maximum

References

1. Woolf SH (1993) Practice guidelines: a new reality in medicine. III. Impact on patient care. Arch Intern Med 153:2646–2655. doi:10.1001/archinte.1993.00410230060008
2. Davis D (2006) Assessing and scaling the knowledge pyramid: the good-guideline guide. Can Med Assoc J 174:337–338. doi:10.1503/cmaj.051611
3. Berlin JA, Golub RM (2014) Meta-analysis as evidence: building a better pyramid. JAMA 312:603. doi:10.1001/jama.2014.8167
4. Brouwers MC, Kho ME, Browman GP et al (2010) AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ Can Med Assoc J 182:E839–E842. doi:10.1503/cmaj.090449
5. McMahon AD (2002) Study control, violators, inclusion criteria and defining explanatory and pragmatic trials. Stat Med 21:1365–1376. doi:10.1002/sim.1120

6. Mehta RH, Chen AY, Alexander KP et al (2015) Doing the right things and doing them the right way: association between hospital guideline adherence, dosing safety, and outcomes among patients with acute coronary syndrome. Circulation 131:980–987. doi:10.1161/CIRCULATIONAHA.114.013451
7. Dawes AJ, Sacks GD, Cryer HG et al (2015) Compliance with evidence-based guidelines and interhospital variation in mortality for patients with severe traumatic brain injury. JAMA Surg 150:965–972. doi:10.1001/jamasurg.2015.1678
8. Bristow RE, Chang J, Ziogas A et al (2015) Impact of national cancer institute comprehensive cancer centers on ovarian cancer treatment and survival. J Am Coll Surg 220:940–950. doi:10.1016/j.jamcollsurg.2015.01.056
9. Schroeck FR, Kaufman SR, Jacobs BL et al (2014) Adherence to performance measures and outcomes among men treated for prostate cancer. J Urol 192:743–748. doi:10.1016/j.juro.2014.03.091
10. Boland GM, Chang GJ, Haynes AB et al (2013) Association between adherence to national comprehensive cancer network treatment guidelines and improved survival in patients with colon cancer: guideline treatment improves survival. Cancer 119:1593–1601. doi:10.1002/cncr.27935
11. van de Water W, Bastiaannet E, Dekkers OM et al (2012) Adherence to treatment guidelines and survival in patients with early-stage breast cancer by age at diagnosis. Br J Surg 99:813–820. doi:10.1002/bjs.8743
12. Feinstein AR (1985) Clinical epidemiology: the architecture of clinical research. Saunders, Philadelphia
13. Cheng SH, Wang CJ, Lin J-L et al (2009) Adherence to quality indicators and survival in patients with breast cancer. Med Care 47:217–225. doi:10.1097/MLR.0b013e3181893c4a
14. Hernan MA (2006) Estimating causal effects from epidemiological data. J Epidemiol Community Health 60:578–586. doi:10.1136/jech.2004.029496
15. Williamson EJ, Aitken Z, Lawrie J et al (2014) Introduction to causal diagrams for confounder selection. Respirology 19:303–311. doi:10.1111/resp.12238
16. Zagouri F, Liakou P, Bartsch R et al (2015) Discrepancies between ESMO and NCCN breast cancer guidelines: an appraisal. The Breast 24:513–523. doi:10.1016/j.breast.2015.02.031
17. National Collaborating Centre for Cancer (2009) Early and locally advanced breast cancer diagnosis and treatment: full guideline. National Collaborating Centre for Cancer, Cardiff https://www.nice.org.uk/guidance/cg80/resources/early-and-locally-advanced-breast-cancer-diagnosis-and-treatment-975682170565. Accessed 12 Dec 2016
18. Senkus E, Kyriakides S, Ohno S et al (2015) Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 26:v8–v30. doi:10.1093/annonc/mdv298
19. Bleicher RJ, Ruth K, Sigurdson ER et al (2016) Time to Surgery and Breast Cancer Survival in the United States. JAMA Oncol 2:330–339. doi:10.1001/jamaoncol.2015.4508
20. Meiers P, Cil T, Guller U, Zuber M (2013) Sentinel lymph node biopsy in early-stage breast cancer patients: improved survival through better staging? Langenbecks Arch Surg Dtsch Ges Für Chir 398:687–690. doi:10.1007/s00423-012-1037-2
21. Coates AS, Winer EP, Goldhirsch A et al (2015) Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. Ann Oncol 26:1533–1546. doi:10.1093/annonc/mdv221
22. Biglia N, Ponzone R, Bounous VE et al (2014) Role of re-excision for positive and close resection margins in patients treated with breast-conserving surgery. Breast Edinb Scotl 23:870–875. doi:10.1016/j.breast.2014.09.009
23. A'Hern RP, Ebbs SR, Baum MB (1988) Does chemotherapy improve survival in advanced breast cancer? A statistical overview. Br J Cancer 57:615–618
24. Bergh J, Jönsson PE, Glimelius B et al (2001) A systematic overview of chemotherapy effects in breast cancer. Acta Oncol Stockh Swed 40:253–281. doi:10.1080/02841860120784
25. Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (2011) Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. The Lancet 378:771–784. doi:10.1016/S0140-6736(11)60993-8
26. Clarke M, Collins R, Darby S et al (2005) Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. Lancet 366:2087–2106. doi:10.1016/S0140-6736(05)67887-7
27. McGuire SE, Gonzalez-Angulo AM, Huang EH et al (2007) Postmastectomy radiation improves the outcome of patients with locally advanced breast cancer who achieve a pathologic complete response to neoadjuvant chemotherapy. Int J Radiat Oncol Biol Phys 68:1004–1009. doi:10.1016/j.ijrobp.2007.01.023
28. Grunfeld E, Noorani H, McGahan L et al (2002) Surveillance mammography after treatment of primary breast cancer: a systematic review. Breast Edinb Scotl 11:228–235. doi:10.1054/brst.2001.0404
29. Montgomery DA, Krupa K, Cooke TG (2007) Follow-up in breast cancer: does routine clinical examination improve outcome? A systematic review of the literature. Br J Cancer 97:1632–1641. doi:10.1038/sj.bjc.6604065
30. Greenland S, Pearl J, Robins JM (1999) Causal diagrams for epidemiologic research. Epidemiol Camb Mass 10:37–48
31. Textor J, Hardt J, Knüppel S (2011) DAGitty: a graphical tool for analyzing causal diagrams. Epidemiol Camb Mass 22:745. doi:10.1097/EDE.0b013e318225c2be
32. Andreano A, Anghinoni E, Autelitano M et al (2016) Indicators based on registers and administrative data for breast cancer: routine evaluation of oncologic care pathway can be implemented. J Eval Clin Pract 22:62–70. doi:10.1111/jep.12436
33. Quan H, Li B, Couris CM et al (2011) Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. Am J Epidemiol 173:676–682. doi:10.1093/aje/kwq433
34. Deyo RA, Cherkin DC, Ciol MA (1992) Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol 45:613–619. doi:10.1016/0895-4356(92)90133-8
35. Caranci N, Biggeri A, Grisotto L et al (2010) L'indice di deprivazione italiano a livello di sezione di censimento: definizione, descrizione e associazione con la mortalità. Epidemiol Prev 34:167
36. Nuova Anagrafe Regionale. http://basidati.agid.gov.it/catalogo/bd?code=154911. Accessed 12 Dec 2016
37. Hernán M, Robins JM (2017) Causal survival analysis. In: Causal Inference. Chapman & Hall/CRC, Boca Raton, partII, pp 75–77
38. van Buuren S (2007) Multiple imputation of discrete and continuous data by fully conditional specification. Stat Method Med Res 16:219–242. doi:10.1177/0962280206074463
39. Rubin DB (2004) Multiple imputation for nonresponse in surveys. John Wiley & Sons, Hoboken
40. SAS Institute Inc. (2013) SAS/STAT Software, Version 9.4., Cary, NC
41. Core Team R (2015) R: A Language and Environment for Statistical Computing. R Found Stat Computing, Vienna, Austria
42. Steuten L, Vrijhoef B, Severens H et al (2006) Are we measuring what matters in health technology assessment of disease management? Systematic literature review. Int J Technol Assess Health Care 22:47–57. doi:10.1017/S0266462306050835

43. Jacobs R, Goddard M (2007) How do performance indicators add up? An examination of composite indicators in public services. Public Money Manag 27:103–110. doi:10.1111/j.1467-9302.2007.00565.x
44. Shwartz M, Restuccia JD, Rosen AK (2015) Composite measures of health care provider performance: a description of approaches: composite measures of health care provider performance. Milbank Q 93:788–825. doi:10.1111/1468-0009.12165
45. Reeves D, Campbell SM, Adams J et al (2007) Combining multiple indicators of clinical quality. Med Care 45:489–496. doi:10.1097/MLR.0b013e31803bb479
46. Huisman M, Kunst A, Andersen O et al (2004) Socioeconomic inequalities in mortality among elderly people in 11 European populations. J Epidemiol Community Health 58:468–475. doi:10.1136/jech.2003.010496