BMJ Open

Use of real-world data to study health services utilisation and comorbidities in long-term breast cancer survivors (the SURBCAN study): study protocol for a longitudinal population-based cohort study

Anna Jansana [1,2] Isabel Del Cura [2,3] Alexandra Prados-Torres,[2,4]
Teresa Sanz Cuesta [2,5] Beatriz Poblador-Plou,[2,4] A Gimeno Miguel [2,4]
Manuela Lanzuela,[6] Berta Ibañez,[2,7] Ibai Tamayo,[2,7] Conchi Moreno-Iribas,[2,7]
María Padilla-Ruiz,[2,8] Maximino Redondo,[2,8] Mercè Comas,[1,2] Laia Domingo,[1,2]
Antonio Díaz-Holgado,[9] Francisco Javier Salamanca,[10] Xavier Castells,[1,2]
Maria Sala [1,2] The SURBCAN group

For numbered affiliations see end of article.

Correspondence to
Dr Maria Sala;
Msalaserra@parcdesalutmar.cat

ABSTRACT

Introduction Breast cancer has become a chronic disease due to survival improvement and the need to monitor the side effects of treatment and the disease itself. The aim of the SURBCAN study is to describe comorbidity, healthcare services use and adherence to preventive recommendations in long-term breast cancer survivors and to compare them with those in women without this diagnosis in order to improve and adapt the care response to this group of survivors.

Methods and analysis Population-based retrospective cohort study using real-world data from cancer registries and linked electronic medical records in five Spanish regions. Long-term breast cancer survivors diagnosed between 2000 and 2006 will be identified and matched by age and administrative health area with women without this diagnosis. Sociodemographic and clinical variables including comorbidities and variables on the use of health services between 2012 and 2016 will be obtained from databases in primary and hospital care. Health services use will be assessed through the annual number of visits to primary care professionals and to specialists and through annual imaging and laboratory tests. Factors associated with healthcare utilisation and comorbidities will be analysed using multilevel logistic regression models. Recruitment started in December 2018.

Ethics and dissemination This study was approved by the Ethics Committee of Parc de Salut Mar. The results of the study will be published in a peer-reviewed journal and will be presented at national and international scientific conferences and at patient associations.

Trial registration number This protocol is registered in Clinical Trials.gov (identifier: NCT03846999).

Strengths and limitations of this study

► To our knowledge, this is the first study focusing on long-term breast cancer survivors and aiming to assess survivors' use of health services according to their health status and comorbidities. It will also assess adherence to the main preventive recommendations.
► In addition, this study is based on real-world data, which allows patient behaviour to be tracked in terms of health services use according to their health status and socioeconomic situation during the 5-year follow-up.
► The study design includes a group of women who have survived at least 5 years since the breast cancer diagnosis and a second group of women without a history of breast cancer which allows to compare the healthcare use between both groups.
► A limitation of the study may be the differences in the availability and quality of the data in the participating regions.
► Our findings will be valuable to improve current practices, estimate the resources needed and to assess adherence to follow-up recommendations.

INTRODUCTION

The term long-term breast cancer survivor (BCS) defines women surviving at least 5 years free from cancer recurrence or new primary cancer.[1] The number of BCS is gradually increasing, particularly in Western countries, and trends suggest their number will continue to increase due to improvements in breast cancer screening, advances in diagnosis and

treatment and greater life expectancy.[2 3] Recent data from the National Cancer Institute show that in 2019, when all stages of cancer are combined, the overall 5-year relative breast cancer survival rate is 90%.[4–6] In Europe, the age-standardised 5-year survival rates are around 87% in the Nordic countries and 83.7% in Spain.[7 8] This increase has led to a growing need by clinicians and researchers to study and assess the needs and patient-centred outcomes of long-term BCS.[3 9]

Research on healthcare services use and comorbidities in long-term BCS has been established as a priority by The National Cancer Institute of the USA, which has included the assessment of the needs and challenges of a growing population of cancer survivors among their six priorities for cancer-related public health research.[10] Cancer survivors' follow-up is a phase of cancer care and includes surveillance, symptom control, treatment of comorbidities and complications, psychosocial needs and care coordination mechanisms. Several European organisations, such as the National Health Service in the UK, the European Breast Cancer Coalition and the Spanish Medical Oncology Society (SEOM), have published recent recommendations regarding the characteristics of follow-up and recommendations on the tests that should be provided to this group of women. All of them highlight the importance of providing a collaborative and coordinated follow-up shared between primary and specialised care.[11–13] These guidelines also include recommendations on the assessment and management of comorbidities in long-term BCS.

Comorbidities associated with both the side effects of breast cancer treatment and with ageing have been widely described, although research has focused on survivors overall rather than on long-term BCS specifically.[14–20] In 2016, the International Consortium for Health Outcomes Measurement (ICHOM) described the main comorbidities associated with breast cancer, which include a wide range of patient-related outcomes as well as clinical health disorders such as arthralgia, anxiety and cardiovascular disorders.[21] Bradshaw et al[22] showed that long-term BCS are at higher risk for cardiovascular disease-related mortality than women without breast cancer. Moreover, 30% of BCS reported symptoms of mental distress during the survival period[23] and around 30% had hypertension and hypercholesterolemia.[24 25]

In contrast, very few studies have so far evaluated health services use and adherence to recommendations for long-term BCS. The factors reported to be associated with lower adherence include higher age at diagnosis, the presence of comorbidities, and a low educational level,[26 27] but there is little information on long-term survivors.[28–30] Another important factor that can also affect adherence is which recommendations are made to patients by health professionals and how they are delivered.

Answers to public health questions can be provided by population cohorts based on real-world data from electronic medical records (EMR) and clinical-administrative data and take into account participants' global context.

Such studies are the preferred options in research in health services.[31 32] Nevertheless, to our knowledge, currently, there is no cohort of long-term BCS using EMR either in Spain or elsewhere in southern Europe.[33 34]

In 2016 the SURBCAN study (SUrvival Breast CANcer Cohort) was initiated by the REDISSEC (*Spanish acronym for: Red de Investigación en Servicios de Salud en Enfermedades Crónicas* (Health Services Research on Chronic Patients Network)). This project combines the experience of expert groups in longitudinal analysis, health services assessment, analysis of comorbidities and chronic diseases belonging to the REDISSEC network. REDISSEC conducts health services research projects to provide solid evidence of best policies, practices and organisational models to address chronicity.[35]

The SURBCAN cohort includes long-term BCS as well as controls with no a history of breast cancer, matched by age and administrative health area treated in the Spanish National Health System (SNS), in order to characterise the care provided by the SNS to long-term BCS and to compare their health outcomes. The SNS is based on the right of every citizen to health protection, public financing and provision of services that ensure the quality of care, among other rights. The SNS is decentralised in autonomous communities (AC). The decentralisation process, finalised in 2001, aimed to adapt health management to the territorial and demographic characteristics of each AC. Furthermore, the Ministry of Health created a Health Information System within the SNS to ensure the availability of information and communication between the government and the AC.

METHODS AND ANALYSIS
Primary objective
The main aim of this cohort is to describe comorbidity, healthcare services use and adherence to preventive recommendations among women surviving breast cancer for 5 years or more.

Design and setting
The SURBCAN study is an observational population-based retrospective cohort study that includes long-term BCS and a control group of women without history of breast cancer from five Spanish areas: Hospital del Mar and SIDIAP (Information System for Primary Care Research in Catalonia), Hospital Costa del Sol in Andalusia, Hospital 12 de Octubre and Primary Care in Madrid, Navarre and the EpiChron Cohort in Aragón. This study protocol was developed on the basis of the Strengthening the Reporting of Observational Studies in Epidemiology checklist statement for cohort studies (online supplementary material 1).[36 37] The recruitment started at the end of 2018 after having obtained the ethical approval.

Study population
All women≥18 years old with a diagnosis of incident breast cancer between 2000 and 2006 and surviving at
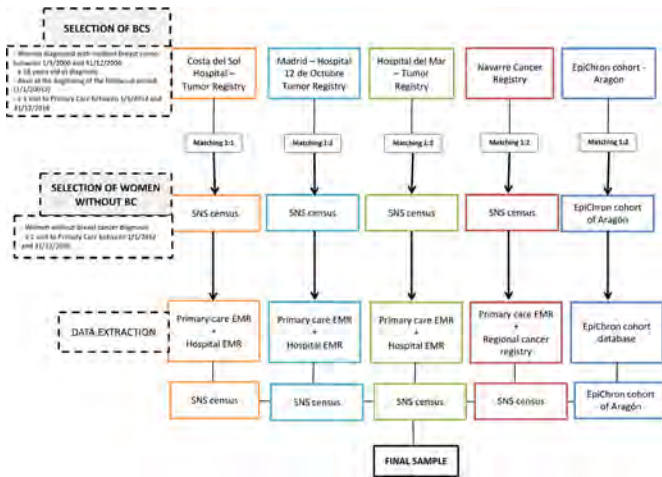
Figure 1 Flowchart of the SURBCAN study. BCS, breast cancer survivor; SNS, Sistema nacional de salud (National Health Service); EMR, electronic medical record.

least 5 years after diagnosis will be identified in each of the subcohorts participating in this study. To ensure a minimum 5-year survival period, all included women must be alive at the beginning of the survival period (1 January 2012). Women with unknown date of diagnosis will be excluded.

In addition, BCS will be only included if they will have attended at least one visit to primary care health services during the follow-up period (01 January 2012 to 31 December 2016). The SURBCAN cohort also includes a control group composed of selected women without a breast cancer diagnosis, matched by age and administrative health area, with BCS. Women in the control group will be identified through the primary care registries of each participating area. All data contained in the cohort will be anonymised. Data collection in the participating centres is described in figure 1.

Information source and data collection

The SURBCAN cohort is based on real-world data from EMR (includes primary and hospital care) and from tumour registries (hospital and population-based registries). The variables collected in the SURBCAN cohort are drawn from patients' routine contacts with the taxpayer-funded health system in Spain (figure 2). Sociodemographic and lifestyle information, pharmaceutical prescriptions, comorbidities and all visits (including emergency visits and hospital admissions) and tests performed in both primary and specialised care will be obtained through databases from the SIDIAP, the Health Service of Madrid, the Health Service of Andalusia, Osansunbidea (Navarre Health Service) and the EpiChron Cohort from Aragón for both BCS and controls. Information on tumour characteristics and cancer treatment will be obtained from hospital and regional tumour registries for BCS. Sociodemographic and lifestyle information will be obtained in the first contact with the health system in the follow-up period (1 January 2012 to 31 December 2016) for each woman. Information is obtained on each visit to primary, specialised and hospital care (date of visit, type of visit and professional visited), all diagnostic tests (date and type of test) and referrals for the whole follow-up period in order to obtain the history of contacts to healthcare services. Comorbidities and drug prescriptions will be extracted at the beginning (first contact) and at the end (last contact) of follow-up.

To analyse comorbidities, the original diagnoses from the International Disease Classification (IDC-9, IDC-10) and from the International Classification of Primary Care will be coded according to the expanded diagnosis clusters (EDCs) using the list of 114 chronic EDCs developed by Salisbury. This list groups together diseases that describe similar or related conditions and is internationally used to define comorbidity and multimorbidity.[38 39] In addition, comorbidities derived from breast cancer as defined by ICHOM are also included. Diagnoses will be obtained from the Primary Care Clinical Database (PCCD) of each region included in the study.[40]

Drugs will be coded using the International Anatomical Therapeutic Chemical classification for medicines, maintained by the WHO. These women will be included in the pertinent analysis until the date of withdrawal.

Sample size calculation

The sample size is estimated on the comparison of two means of the number of visits to hospital outpatient consultations per woman, matched per case and two controls. A female BCS has an average of 18.4 consultations (preliminary data from Hospital del Mar in 2012) and we aimed to detect a difference with women without breast cancer of at least three consultations. Assuming a SD of 25, a significance level of 5% and a statistical power of 80%, we needed a sample of 1619 cases and 3238 controls. The Hospital del Mar tumour registry, which covers a population of 300 000 inhabitants, contained data on 1676 women with breast cancer diagnosed and treated in the same hospital between 2005 and 2010 and who survived a minimum of 5 years after diagnosis.

Given that this sample already guaranteed the minimum number of cases for the study, the inclusion of cases from
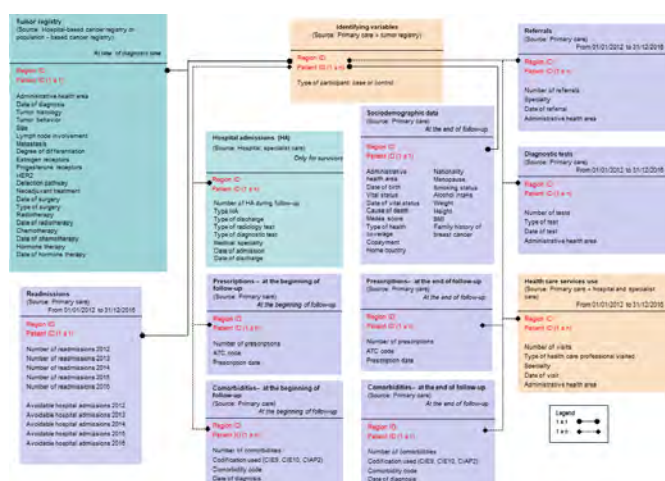

Figure 2 Variables included in the SURBCAN study.

the rest of the participating areas will allow all the study questions set to be answered and the necessary adjustments to be made.

Statistical analysis

Demographic and clinical characteristics of the patients will be described, by group (BCS/control group) and by area, using mean and SD, median and interquartile ranges, or frequencies with proportions, depending on the nature of the variables. The pattern of health services utilisation will be described through rates per 100 person-years. Differences in demographic and clinical variables between BCS and the control group will be assessed by parametric and non-parametric tests, such as the Student's *t*-test, or the Mann-Whitney *U*-test, whereas differences in utilisation rates will be assessed using incidence rate ratios, with 95% CIs from the exactci package in R.

For the analyses of comorbidities, diagnoses associated with chronic diseases will be identified to differentiate them from acute diagnoses. The prevalence of comorbidities will be compared between groups and according to certain sociodemographic variables (age group, subcohorts, nationality, administrative health area and deprivation index) using the χ^2 test, the trend for proportions test or other parametric and non-parametric tests, depending on the nature of the variables. The presence of specific comorbidities by group (and also according to the characteristics of the primary tumour and the treatment received by the BCS group) will also be assessed with logistic regression models, to account for possible confounding variables, which will provide ORs with 95% CIs. The incidence of new comorbidities in the follow-up will also be modelled with the same methodology.

For the analysis of health services utilisation, generalised linear models that may have different specifications depending on the distribution of the dependent variables (number of visits, number of hospital admissions, etc) will be fitted. For rare events, such as the number of hospital admissions, count models (Poisson, negative binomial or zero-inflated models) will be fitted, including group as exposure, exposure-time (total follow-up time, taking into account possible censorship such as death or loss to follow-up due to other causes) as offset and clinical and demographic characteristics as potentially confounding variables. For more frequent data, such as the number of visits, linear models will be fitted, with log-transformed data if necessary and the same covariates as for count data. Complementarily, a sensitivity analysis will be performed to determine which factors have the strongest influence on the variation in resource consumption.

All analyses, both crude and adjusted, will be stratified by BCS and by women with no history of breast cancer in order to study the distribution of confounders in the sample. As this is a matched cohort by age and administrative health area, a similar distribution of sociodemographic is expected.

Additionally, the level of adherence to the American Society of Clinical Oncology and SEOM breast cancer follow-up guidelines during the study period according to tumour characteristics and time from diagnosis (5–10 years and 10 or more years) will be analysed using logistic regression, considering as dependent variables whether or not each recommendation was fulfilled.

Regarding missing data, women with unknown information in the variables of interest will be excluded from each analysis. Each analysis will include in its description the number of missing and it will be decided to impute or not according to this number. However, sensitivity analyses will be carried out to study the effect of missing data.

Quality assurance

The main challenge for the construction of a common and systematic clinical database is the need to bring together the existing decentralised information in the different medical records of the AC.[39] All the team members and the data managers in charge of each subcohort will be are provided with a handbook drafted by the main researcher's team, which includes the overall description of the project, the team members and their contacts, the structure and methodology, and a detailed definition of all the variables included in the project. In addition, they will be provided with a quality control protocol for the SURBCAN database, which defines the accepted values for each variable, the range of possible values, and logical validation rules for the entire study population. Each subcohort will carry out case identification in their area and anonymisation following the steps described in the protocols.

The project has a web platform that works as a data deployment web. It has private access with a password. The data manager of each subcohort will upload data through this platform. By means of error detection software, the platform identifies the mistakes or inconsistencies in the uploaded data compared with specifications in the quality control protocol, so that data managers can correct and ensure data quality and homogeneity between the subcohorts. Once all the data have been uploaded, the entire database will be validated by the main researcher's team, who will also merge the databases, ensuring that the final database follows standards of good practice. During that process, a monthly online conference will be held to review the ongoing process and supervise the coordination of the study.

Patient and public involvement

The purpose of this study derives from previous research from our research group where the need arose to study the health needs of long-term BCSs and their relationship with health services. In previous studies, we have been in contact with patient associations as we believe it is essential to include their vision and current opinions.

Although patients and the public were neither involved in developing the hypothesis, the specific aims or the research question, nor were they involved in developing a plan for design or implementation of the study, the

results and conclusions will be discussed with members of patient with breast cancer associations in Spain.

DISCUSSION

The SURBCAN study provides a novel approach to describe comorbidity patterns and to assess healthcare services use based on real-world data and focused on long-term BCS.

The analysis of the data provided by this study will be used to assess the peculiarities of the care provided to long-term BCS in the Spanish National Health Service of Spain. The results of this study will allow estimation of the type of healthcare used by these women according to their comorbidities and/or level of risk, and their resource consumption compared with that of women without a history of breast cancer.

The conclusions and outcomes of this study will be useful for multiple stakeholders, including decision-makers, healthcare providers, medical organisations and patients to guide their decisions and to understand the most appropriate care for each BCS. This will be valuable information to improve current practices and health outcomes, to estimate the resources needed to deal with the increase in BCS in coming years and to assess adherence to follow-up recommendations.

Strengths and weaknesses

The main strength of the SURBCAN study is that it is based on real-world data, which enables us to track patient behaviour regarding health services according to their health status and socioeconomic situation. Real-world data were obtained through EMR databases that include all the information from all levels of care of the Spanish health system: primary care, hospital care, specialised care, emergency room visits and consumption of resources, such as diagnostic and laboratory tests, as well as drug prescriptions.

EMR were created with a clear healthcare purpose and provide large sample sizes, broad representativeness of participants, minimisation of memory bias and an almost complete medical history including comorbidity information. EMR allow chronic processes to be analysed from a multicausal approach as they include data from diverse sources. The quality of these databases has also been validated for the study of various chronic diseases.[41 42]

However, the availability of data can differ in each subcohort, since although all of them are within the framework of the Spanish National Health Service, each region has a different health service organisation and management models. For instance, date of diagnosis of comorbidities present at the first contact will not be available in all five regions given that EMR started in different years: in Hospital del Mar–SIDIAP the EMR started in 2005, in Hospital Costa del Sol in 2007, in Hospital 12 de Octubre–Primary care Service of Madrid in 2003, in Navarre the EMR of Primary Care started in 2004 and the EMR of specialised care in 2007 and in Aragón the EMR

started in 2006. Detailed information regarding tumour characteristics and cancer treatment may not be fully available. To remedy the lack of information, depending on the number of missing, specific statistical analyses will be carried out. However, it is important to note that the methodology of real-world data analysis includes a less conservative treatment of unknown data. This translates into analysing the data as objectively as possible in order to understand the reality of the health system.

ETHICS AND DISSEMINATION

This project will be carried out in accordance with the basic principles of protection of human rights and dignity, as stated in the Declaration of Helsinki and in accordance with the regulations in force. The Hospital del Mar Medical Research Institute, as the coordinating centre, obtained approval from its ethics committee. This study was approved by the Ethics Committee of Parc de Salut Mar (Catalonia), the Clinical Research Ethics Committee of the Government of Navarre, the Ethics Committee of the Jordi Gol University Institute for Research in Primary Care (Catalonia), the Ethics Committee of the Costa del Sol Health Agency (Andalusia), the Central Research Commission of the Community of Madrid and the Ethics Committee of Aragón.

All information obtained will be treated confidentially and anonymously, in compliance with Organic Law 3/2018 of 6 December. The usual care and attendance of the study population were unaffected and there were no changes or actions on treatments or visits, or collection or analysis of biological samples. Informed consent from patients was not necessary as the SURBCAN study is a retrospective observational project that only includes anonymised data from clinical and administrative databases. The findings of this study will be published in a peer-reviewed journal and presented at national and international scientific conferences in order to disseminate the results to academic, health professional and patient associations. The results will be available on our website to participants and to a wider public at the time of publication.

Author affiliations
[1]Department of Epidemiology and Evaluation, Hospital del Mar Institute for Medical Research, Barcelona, Barcelona, Spain
[2]Health Services Research on Chronic Patients Network (REDISSEC), Instituto de Salud Carlos III, Madrid, Madrid, Spain
[3]Primary Care Research Unit, Madrid Health Service, Madrid, Madrid, Spain
[4]EpiChron Research Group on Chronic Diseases, IACS, IIS Aragon, Miguel Servet University Hospital, Zaragoza, Zaragoza, Spain
[5]Primary Care Research Unit, Comunidad de Madrid Servicio Madrileno de Salud, Madrid, Spain
[6]Radiotherapy Department, Miguel Servet University Hospital, Zaragoza, Aragón, Spain
[7]Unit of Methodology, Navarrabiomed-Fundación Miguel Servet, Pamplona, Navarra, Spain
[8]Research Unit - IBIMA, Hospital Costa del Sol, Marbella, Andalucía, Spain
[9]Information System Unit, Directorate for Public Health, Health Service of Madrid, Madrid, Spain
[10]Tumor registry, Hospital Universitario Doce de Octubre, Madrid, Madrid, Spain

ORCID iDs
Anna Jansana http://orcid.org/0000-0001-8504-1618
Isabel Del Cura http://orcid.org/0000-0002-3931-5304
Teresa Sanz Cuesta http://orcid.org/0000-0001-8791-8030
A Gimeno Miguel http://orcid.org/0000-0002-5440-1710
Maria Sala http://orcid.org/0000-0002-9955-8746

REFERENCES

1. Cardoso F, Harbeck N, Barrios CH, et al. Research needs in breast cancer. *Ann Oncol* 2017;28:208–17.
2. Boyle P, Levin B. *World cancer report 2008*. Geneva: International Agency for Research on Cancer (IARC), 2010.
3. Erikson C, Salsberg E, Forte G, et al. Future simply and demand for oncologist: challenges to assuring access to oncology services. *J OncolPract* 2007;3:79–86.
4. Noone A, Howlader N, Krapcho M, et al. *SEER cancer statistics review, 1975-2015*. Bethesda: National Cancer Institute, 2015. https://seer.cancer.gov/csr/1975_2015/
5. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
6. SEERStat. Surveillance, epidemiology, and end results (seer) program. SEERStat database: Incidence-SEER-9 regulations. National cancer Institute, division of cancer control and population sciences, 2003. Available: https://seer.cancer.gov/ [Accessed 20 Jul 2018].
7. Baeyens-Fernández JA, Molina-Portillo E, Pollán M, et al. Trends in incidence, mortality and survival in women with breast cancer from 1985 to 2012 in Granada, Spain: a population-based study. *BMC Cancer* 2018;18:781.
8. Engholm G, Ferlay J, Christensen N, et al. NORDCAN--a Nordic tool for cancer information, planning, quality control and research. *Acta Oncol* 2010;49:725–36.
9. Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin* 2016;66:271–89.
10. National Cancer Institute. Public health research and cancer. Available: https://www.cancer.gov/research/areas/public-health#opportunities [Accessed 25 Jan 2019].
11. Runowicz CD, Leach CR, Henry NL, et al. American cancer Society/American Society of clinical oncology breast cancer survivorship care guideline. *CA Cancer J Clin* 2016;66:43–73.
12. Alba E, Cruz J, Barón F, et al. Monográfico SEOM de largos supervivientes en cáncer. Sociedad Española de Oncología Médica, 2012. Available: https://seom.org/seomcms/images/stories/recursos/infopublico/publicaciones/MONOGRAFICO_SEOM_LARGOS_SUPERVIVIENTES_I.pdf
13. Hewitt M, Greenfield S, Stovall E. Delivering cancer survivorship care – from cancer patient to cancer survivor: lost in transition. *TheNationalAcademiesPress. Washington DC* 2006:187–321.
14. Ewertz M, Land LH, Dalton SO, et al. Influence of specific comorbidities on survival after early-stage breast cancer. *Acta Oncol* 2018;57:129–34.
15. Jordan JH, Thwin SS, Lash TL, et al. Incident comorbidities and all-cause mortality among 5-year survivors of stage I and II breast cancer diagnosed at age 65 or older: a prospective-matched cohort study. *Breast Cancer Res Treat* 2014;146:401–9.
16. Snyder CF, Frick KD, Herbert RJ, et al. Quality of care for comorbid conditions during the transition to survivorship: differences between cancer survivors and noncancer controls. *J Clin Oncol* 2013;31:1140–8.
17. Edwards MJ, Campbell ID, Lawrenson RA, et al. Influence of comorbidity on chemotherapy use for early breast cancer: systematic review and meta-analysis. *Breast Cancer Res Treat* 2017;165:17–39.
18. Parés-Badell O, Banqué M, Macià F, et al. Impact of comorbidity on survival by tumour location: breast, colorectal and lung cancer (2000-2014). *Cancer Epidemiol* 2017;49:66–74.
19. Leach CR, Weaver KE, Aziz NM, et al. The complex health profile of long-term cancer survivors: prevalence and predictors of comorbid conditions. *J Cancer Surviv* 2015;9:239–51.
20. Mandelblatt JS, Lawrence WF, Cullen J, et al. Patterns of care in early-stage breast cancer survivors in the first year after cessation of active treatment. *J Clin Oncol* 2006;24:77–84.
21. Ong WL, Schouwenburg MG, van Bommel ACM, et al. A standard set of value-based patient-centered outcomes for breast cancer: the International Consortium for health outcomes measurement (ICHOM) initiative. *JAMA Oncol* 2017;3:677–85.
22. Bradshaw PT, Stevens J, Khankari N, et al. Cardiovascular disease mortality among breast cancer survivors. *Epidemiology* 2016;27:6–13.
23. Syrowatka A, Motulsky A, Kurteva S, et al. Predictors of distress in female breast cancer survivors: a systematic review. *Breast Cancer Res Treat* 2017;165:229–45.
24. Nelson SH, Marinac CR, Patterson RE, et al. Impact of very low physical activity, BMI, and comorbidities on mortality among breast cancer survivors. *Breast Cancer Res Treat* 2016;155:551–7.
25. Leysen L, Beckwée D, Nijs J, et al. Risk factors of pain in breast cancer survivors: a systematic review and meta-analysis. *Support Care Cancer* 2017;25:3607–43.
26. Jansana A, Posso M, Guerrero I, et al. Health care services use among long-term breast cancer survivors: a systematic review. *J Cancer Surviv* 2019;13:477–93.
27. Patnaik JL, Byers T, Diguiseppi C, et al. The influence of comorbidities on overall survival among older women diagnosed with breast cancer. *J Natl Cancer Inst* 2011;103:1101–11.
28. Schneble EJ, Graham LJ, Shupe MP, et al. Current approaches and challenges in early detection of breast cancer recurrence. *J Cancer* 2014;5:281–90.
29. Khatcheressian JL, Hurley P, Bantug E, et al. Breast cancer follow-up and management after primary treatment: American Society of clinical oncology clinical practice guideline update. *J Clin Oncol* 2013;31:961–5.
30. Yeom YK, Chae EY, Kim HH, et al. Screening mammography for second breast cancers in women with history of early-stage breast cancer: factors and causes associated with non-detection. *BMC Med Imaging* 2019;19:2.

31 Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. *Annu Rev Public Health* 2011;32:91–108.

32 Prados-Torres A, Poblador-Plou B, Gimeno-Miguel A, *et al*. Cohort profile: the epidemiology of chronic diseases and multimorbidity. The EpiChron cohort study. *Int J Epidemiol* 2018;47:382–4.

33 McBride ML, Groome PA, Decker K, *et al*. Adherence to quality breast cancer survivorship care in four Canadian provinces: a CanIMPACT retrospective cohort study. *BMC Cancer* 2019;19:659.

34 Grandjean I, Kwast ABG, de Vries H, KwastAB deVH, Klaase J, *et al*. Evaluation of the adherence to follow-up care guidelines for women with breast cancer. *Eur J Oncol Nurs* 2012;16:281–5.

35 REDISSEC. Red de Investigación en servicios de salud en Enfermedades crónicas, 2013. Available: https://www.redissec.com/es/ [Accessed 20 Jul 2019].

36 Benchimol EI, Smeeth L, Guttmann A, *et al*. The reporting of studies conducted using observational Routinely-collected health data (record) statement. *PLoS Med* 2015;12:e1001885.

37 Vandenbroucke JP*et al*. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med* 2007;147:W–194.

38 Quan H, Sundararajan V, Halfon P, *et al*. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43:1130–9.

39 Salisbury C, Johnson L, Purdy S, *et al*. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2011;61:e12–21.

40 MSSSI. Base de datos clínicos de atención primaria del sistema nacional de salud-BDCAP, 2014. Available: https://www.mscbs.gob.es/estadEstudios/estadisticas/estadisticas/estMinisterio/SIAP/Diseno_BDCAP.pdf [Accessed 10 Jul 2019].

41 Birkhead GS. Successes and continued challenges of electronic health records for chronic disease surveillance. *Am J Public Health* 2017;107:1365–7.

42 Kiasuwa Mbengi RL, Nicolaie AM, Goetghebeur E, *et al*. Assessing factors associated with long-term work disability after cancer in Belgium: a population-based cohort study using competing risks analysis with a 7-year follow-up. *BMJ Open* 2018;8:e014094.