IQWiG Reports – Commission No. V18-05

Relationship between volume of services and quality of treatment outcome in the surgical treatment of breast cancer - rapid report[1]

Extract

[1] Translation of Chapters 1 to 7 of the rapid report V18-05 *Zusammenhang zwischen Leistungsmenge und Qualität des Behandlungsergebnisses bei der chirurgischen Behandlung des Brustkrebses* (Version 1.0; Status: 13 January 2020 [German original], 3 March 2020 [English translation]). Please note: This document was translated by an external translator and is provided as a service by IQWiG to English-language readers. However, solely the German original text is absolutely authoritative and legally binding.

Publishing details

Publisher:

Institute for Quality and Efficiency in Health Care

Topic:

Relationship between volume of services and quality of treatment outcome in the surgical treatment of breast cancer - rapid report

Commissioning agency:

Federal Joint Committee

Commission awarded on:

20 December 2018

Internal Commission No.:

V18-05

Address of publisher:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln
Germany

Phone: +49 221 35685-0
Fax: +49 221 35685-1
E-mail: berichte@iqwig.de
Internet: www.iqwig.de

This report was prepared in collaboration with external experts.

The responsibility for the contents of the report lies solely with IQWiG.

According to §139b (3) No. 2 of Social Code Book (SGB) V, Statutory Health Insurance, external experts who are involved in the Institute's research commissions must disclose "all connections to interest groups and contract organizations, particularly in the pharmaceutical and medical devices industries, including details on the type and amount of any remuneration received". The Institute received the completed *Form for disclosure of potential conflicts of interest* from each external expert. The information provided was reviewed by a Committee of the Institute specifically established to assess conflicts of interests. The information on conflicts of interest provided by the external experts and external reviewers is presented in Chapter C of the full report. No conflicts of interest were detected that could endanger professional independence with regard to the work on the present commission.

External expert

- Elmar Stickeler, Clinic for Gynaecology and Obstetrics, University Hospital Aachen, Germany

External review of the rapid report

- Bernd Gerber, University Women's Clinic and Polyclinic at the Südstadt Hospital of the Hanseatic City of Rostock, Germany

IQWiG thanks the external experts for their collaboration in the project.

IQWiG employees

- Franziska Vosseberg
- Eva Höfer
- Lutz Altenhofen
- Ralf Bender
- Mandy Kromp
- Matthias Maiworm
- Claudia Mischke
- Dorothea Sow

Keywords: Minimum Volume, Breast Neoplasms, Systematic Review

Key statement

Research question

The aim of this investigation is to

present and assess any correlation between the volume of services (VoS) and the quality of the treatment outcome in the surgical treatment of primary mammary carcinoma (research question 1) and

present studies which investigate the effects of a minimum number of cases introduced into the healthcare system on the quality of the treatment outcome in the surgical treatment of primary mammary carcinoma (research question 2).

If this process reveals data on a correlation between VoS and quality of treatment outcome in palliative surgery, they will be presented as supplementary information.

Conclusion

For the investigation of any correlation between VoS and treatment quality in the surgical treatment of breast cancer, a total of 10 studies were included. One of the 10 included studies was assessed as having a high informative value of results.

For the outcome of all-cause mortality, based on a high informative value of results, a correlation between both hospital and doctor VoS and treatment quality was derived. Hence, a higher VoS is associated with lower mortality rates. For other mortality outcomes (breast cancer mortality and other-cause mortality), based on a low informative value of results, a correlation was found either only on the hospital level or only on the doctor level. Again, the results showed lower mortality for these levels.

With respect to performing a reoperation, based on a low informative value of results, it was possible to derive a correlation between both doctor and hospital VoS and treatment quality. At the hospital level, a correlation between VoS and treatment quality was apparent even in comparison with the intermediate VoS category. The results show lower reoperation rates in higher VoS categories.

No studies of meaningful interpretive value were found which investigated any correlation between VoS and treatment quality reflected by other outcomes, such as adverse effects of therapy, local recurrence, disease-free survival, or health-related quality of life. Further, it was not possible to include studies of meaningful interpretive value investigating any effects of specifically introduced minimum case numbers.

Table of contents

List of tables

Page

List of figures

Page

List of abbreviations

Abbreviation	Meaning
BCS	Breast conserving surgery
DCIS	Ductal carcinoma in situ
DKG	Deutsche Krebsgesellschaft (German Cancer Society)
G-BA	Gemeinsamer Bundesausschuss (Federal Joint Committee)
IQTIG	Institut für Qualität und Transparenz im Gesundheitswesen (Institute for Quality Assurance and Transparency in Health Care)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Institute for Quality and Efficiency in Health Care)
RCT	Randomized controlled trial
RKI	Robert Koch Institute
SGB	Sozialgesetzbuch (Social Code Book)
VoS	Volume of services

1 Background

Correlation between volume of services and quality of treatment outcome

As early as in 1979, Luft et al. examined any correlation between volume of services (VoS) and quality of treatment outcome for 12 surgical procedures of different levels of complexity [1]. Their investigations showed that, for complex surgical procedures, there is a correlation between hospital VoS and the quality of treatment outcome. In the following years, various studies showed a similar correlation for many medical services in different healthcare systems, with the VoS being investigated per hospital and per doctor [2-5].

The legal mandate of the Federal Joint Committee (G-BA) regarding minimum volume rules [6] is based upon the idea that there is a concrete connection between the probability of treatment success and the experience of the parties principally involved in rendering the service [6]. As part of quality assurance of registered hospitals, the G-BA therefore defines a catalogue of plannable services for which the quality of the treatment outcomes is dependent on the VoS provided. This dependency is to be assessed on the basis of appropriate studies [7]. In December 2003, the G-BA for the first time set forth minimum volumes which are binding in Germany in accordance with §137 (3), Sentence 1, No. 2 Social Code Book (SGB) V.

These minimum volume rules are binding for hospitals registered in accordance with §108 SGB V and specify in which cases a hospital may render the services for which minimum volumes have been set forth [8]. Hospitals may render the services in question only if the hospital owner annually declares vis-à-vis the state associations of the statutory health insurers that the specified minimum volume will be met in the next year as well [8]. However, some exceptions apply. For instance, minimum volumes generally do not apply in case of emergency. The state authorities responsible for hospital planning can define exceptions for services where the implementation of minimum volume rules may jeopardize state-wide service provision to the population.

No binding minimum volumes are currently in place for surgical procedures performed to treat breast cancer [8]. However, the German Cancer Society (DKG) demands minimum case numbers as part of its breast centre certification system. Certification as a breast centre requires at least 100 primary cases per centre and 50 breast cancer surgeries per surgeon per year [9]. In 2018, 280 sites were certified as breast centres, and 55 715 primary cases were treated there [10].

Breast cancer

At 69 220 new cases per year, breast cancer is the most common cancer among women in Germany [11]. In rare cases, men can develop breast cancer as well. The German Centre for Cancer Registry Data at the Robert Koch Institute (RKI) recorded 650 new cases in men in 2014. The 10-year relative survival rate in 2014 was 82% for women and 69% for men [11].

In women, age is the main risk factor for the development of mammary carcinoma. Further risk factors that apply exclusively to women are dense breast and glandular tissue, early menarche or late menopause, few or no births, and higher age at the 1st birth. Risk factors affecting both sexes are benign tissue changes and a family history or genetic predisposition. Modifiable risk factors include heavy alcohol consumption, overweight, lack of exercise, and smoking. In men, genetic factors and hormonal changes are the main factors responsible for the development of breast cancer [12-16].

Breast cancer precursors

Breast cancer precursors are cell changes associated with an elevated risk of breast cancer. In addition to lesions of unclear biological potential, these precursors include ductal carcinoma in situ (DCIS), which develops within the milk ducts and has not yet grown into surrounding tissue. Due to the higher risk of it developing into malignant invasive carcinoma when compared to other breast cancer precursors, surgical removal is always recommended for DCIS [13, 14, 17].

Tumour classification

To describe the clinical and pathological spread of a carcinoma, it is classified on the basis of its size (T), lymph node involvement (N), and presence of distant metastases (M) (TNM classification) [14, 17, 18].

Other factors taken into account when classifying the tissue and assessing the chances for recovery include cancer tissue differentiation (grading), proliferation rate (Ki-67 index), and spread in the lymph nodes, the tumour's hormone receptor status, and completeness of resection in case of surgery (R0 resection) [13, 14, 17].

Surgical treatment of mammary carcinoma

In 2014, suspected breast cancer resulted in 74 224 surgical procedures being performed at 817 German hospitals [9]. The surgical treatment of mammary carcinoma can be generally classified as breast-conserving surgery versus mastectomy. The choice of therapeutic procedure depends on the patient's individual physical, psychological, and social situation as well as the age, comorbidities, and preferences of the patient [19].

For men, the surgical treatment of mammary carcinoma is based on the guidelines issued for the surgical treatment of women. However, the preferred surgical technique for men is modified radical mastectomy (mastectomy with lymph node removal) [12].

Breast-conserving surgery

Breast-conserving surgery is now considered the standard for treating both invasive and non-invasive mammary carcinoma and is chosen by about 60–70% of female patients [13]. It takes the form of either tumourectomy, segmentectomy, or quadrantectomy. The goal is complete tumour removal. Breast-conserving therapy with subsequent radiotherapy is equivalent to

mastectomy in terms of survival figures [19-21]. Complete tumour removal (R0/margin-negative resection) is considered important to achieve a low risk of local recurrence [14, 19].

Modified radical mastectomy

Modified radical mastectomy, i.e. complete excision of the mammary gland including the nipple and pectoral fascia, is performed in 20–40% of female patients [13]. Increasingly, skin-sparing forms of mastectomy are being used, either with or without preservation of the nipple-areola complex [14, 19, 22].

Surgical treatment of the axilla

Axillary staging is part of the surgical procedures used to treat invasive mammary carcinoma and serves to determine the histologic nodal status. The procedure is indicated if the lymph node status is found to be negative (cN0) by both palpation and ultrasound. In case of cN0, the current standard is sentinel lymph node biopsy (SLNB) to determine whether a tumour has spread into the lymph vessels. Axillary staging is not indicated in case of clinically suspected positive lymph nodes (cN+) [13, 14, 19]. Rather, axillary dissection is usually indicated in that situation [13, 19].

Palliative treatment

In metastatic mammary carcinoma, palliative therapy aims to alleviate symptoms and increase survival time. It can include surgery, radiotherapy, and drugs. Required measures depend on the patient's individual needs [14, 23].

2 Research question

The aim of this investigation is to

- present and assess any correlation between VoS and the quality of the treatment outcome in the surgical treatment of primary mammary carcinoma (research question 1) and
- present studies which investigate the effects of a minimum number of cases introduced into the healthcare system on the quality of the treatment outcome in the surgical treatment of primary mammary carcinoma (research question 2).

If this process reveals data on a correlation between VoS and quality of treatment outcome in palliative surgery, they will be presented as supplementary information.

3 Course of the project

On 16 August 2018, the Federal Joint Committee (G-BA) commissioned the Institute for Quality and Efficiency in Health Care (IQWiG) with a systematic literature search, including evaluation of the evidence on any correlation between VoS and quality of treatment outcome in the surgical treatment of breast cancer (including palliative surgery). Work on the project started on 15 April 2019.

On the basis of the project outline, a rapid report was generated and additionally subjected to an external review. This report was sent to the G-BA and published 4 weeks later on the IQWiG website.

4 Methods

Due to differences between the research questions, different methods were used in some cases.

4.1 Criteria for study inclusion in the investigation

4.1.1 Population

The assessment included studies with patients who were surgically treated for breast cancer.

4.1.2 VoS

VoS was defined as the number of surgical treatments of breast cancer performed per hospital site, per doctor, or per site-doctor combination within a defined time period.

4.1.3 Outcomes

For the investigation, the following outcomes were examined:

- Mortality, such as
  - overall survival
  - breast cancer mortality
- Morbidity, such as
  - need for re-excision
  - local recurrence
  - disease-free survival
  - adverse effects of therapy, such as
    - serious treatment-related complications
    - serious adverse events
- Health-related quality of life, including activities of daily living and dependence on help from others

The inclusion of any usable data on other outcomes or validated quality indicators, such as lymph node removal in DCIS and in breast-conserving therapy, was also permitted.

4.1.4 Study types

Controlled interventional studies or observational studies (e.g. cohort studies or case control studies) were suitable for both research questions.

For controlled interventional studies, the intervention to be examined was the specification of a minimum number of cases. Possible comparator groups were groups with a different or no specified volume.

4.1.5 Adjustment

In breast cancer surgery, the quality of the treatment outcome is influenced, for instance, by the primary tumour size, axillary lymph node status, grading, and hormone receptor status. Other influencing factors include patient sex, age, and comorbidities as well as the (neo)adjuvant therapy. Further indication-specific risk factors are possible.

Therefore, the control of relevant confounders (risk adjustment) was a prerequisite for study inclusion. Control was assumed to exist if the study analysis involved suitable statistical methods to adjust for relevant confounders in an effort to address the problem of potential structural inequalities (unfair comparisons) between hospitals or treatment providers (doctors, nurses, etc.) with high and low VoS.

Likewise, cluster effects (e.g. greater similarity of outcomes in patients within the same hospital versus patients from different hospitals due to hospital-specific characteristics) had to have been taken into consideration by means of adequate statistical methods.

4.1.6 Study duration

There were no restrictions regarding the study duration.

4.1.7 Publication period

Studies with a publication date of January 2000 or later were included in the study.

4.1.8 Transferability

To ensure the transferability of study results to the German healthcare system, studies from European countries as well as the United States, Canada, Australia, and New Zealand were eligible for inclusion.

For international studies, at least 80% of the data had to come from the above countries.

4.1.9 Tabular presentation of the criteria for study inclusion

Table 1 and Table 2 list the criteria which had to be met by studies to be included in the assessment.

Table 1: Overview of inclusion and exclusion criteria for interventional studies

Inclusion and exclusion criteria	
I1.1	Patients who were surgically treated for breast cancer (also see Section 4.1.1)
I1.2	Study intervention: use of a minimum number of cases (also see Section 4.1.4)
I1.3	Comparator intervention: use of a different or no minimum number of cases (also see Section 4.1.4)
I1.4	Outcomes as formulated in Section 4.1.3
I1.5	Controlled intervention study as formulated in Sections 4.1.4 and 4.1.5
I1.6	Publication date of January 2000 or later
I1.7	Full publication available[a]
I1.8	Studies which are transferable to the German healthcare system (also see Section 4.1.8)
E1.1	Multiple publications without relevant additional information

a: In this context, a study report in accordance with ICH E3 [24] or a report about the study which met the criteria of the TREND statement [25] and allowed an assessment of the study was considered a full publication so long as the information on both the study methods and study results provided in these documents was not confidential.

ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs

Table 2: Overview of inclusion and exclusion criteria for observational studies

Inclusion and exclusion criteria	
I2.1	Patients who were surgically treated for breast cancer (also see Section 4.1.1)
I2.2	Investigation of any correlation between VoS over a certain period and the quality of the treatment outcome (also see Section 4.1.2)
I2.3	Outcomes as formulated in Section 4.1.3
I2.4	Studies as formulated in Section 4.1.4
I2.5	Adjustment as formulated in Section 4.1.5
I2.6	Publication date of January 2000 or later
I2.7	Full publication available[a]
I2.8	Studies which are transferable to the German healthcare system (also see Section 4.1.8)
E2.1	Multiple publications without relevant additional information

a: In this context, a study report in accordance with ICH E3 [24] or a report about the study which met the criteria of the STROBE statement [26] and allowed an assessment of the study was considered a full publication, so long as the information on both the study methods and study results provided in these documents was not confidential.

ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology, VoS: volume of services

4.1.10 Inclusion of studies which do not fully meet the above criteria

In accordance with IQWiG General Methods Version 5.0, Chapter 9 [27], for the inclusion criteria I1.1/I2.1 (population), I1.2 (use of a minimum number of cases), and I1.3 (comparator intervention with respect to the study's comparator group), and/or I2.2 (VoS), and I1.8/I2.8 (transferability), it sufficed if at least 80% of included patients fulfilled these criteria. For such studies, subgroup analyses, if any, on patients who fulfilled the inclusion criteria were used. Studies in which the inclusion criteria I1.1/I2.1, I1.2/I2.2, and I1.3 as well as I1.8/I2.8 were fulfilled by fewer than 80% of patients were included only if subgroup analyses were available for patients who did fulfil the inclusion criteria.

4.2 Comprehensive information retrieval

4.2.1 Sources of information

For the comprehensive information retrieval, a systematic search was conducted for relevant studies or documents in accordance with IQWiG General Methods Version 5.0, Chapter 8 [27]. The following primary and further information sources as well as search techniques were considered:

Primary information sources

- Bibliographic databases
  - MEDLINE
  - Embase
  - Cochrane Central Register of Controlled Trials
  - Cochrane Database of Systematic Reviews

Further information sources and search techniques

- Use of further search techniques
  - Screening of reference lists of systematic reviews found
- Requests to authors

4.2.2 Selection of relevant studies or documents from the results of the bibliographic search

In a first step, the titles and, if available, abstracts of the hits retrieved in the bibliographic databases were screened for potential relevance in terms of the inclusion criteria (see Table 1 and Table 2). In a second step, any documents considered potentially relevant were checked for relevance. Both steps were performed by 2 persons independently of each other. Any discrepancies were resolved by discussion between them.

Selection of relevant studies or documents from further information sources

Search results from the further information sources considered were screened for studies by 1 reviewer. The studies found were then checked for relevance. The whole process was then checked by a 2nd reviewer. Any discrepancies in one of the listed selection steps were resolved by discussion between the 2 reviewers.

4.3 Information synthesis and analysis

4.3.1 Presentation of the individual studies

All information needed for the investigation was extracted from the documents on the included studies and put into standardized tables. Any discrepancies found in connection with the comparison of information from different documents or from multiple data points within the same document, provided such discrepancies had the potential of considerably influencing the interpretation of results, are presented in the results section of the report.

Results were typically omitted from the investigation whenever they were based on fewer than 70% of the patients to be included in the analysis, that is, whenever more than 30% of patients were excluded from analysis.

Results were also omitted from the investigation whenever the percentage of patients excluded from analysis differed by more than 15% between groups.

If the studies' authors used several statistical models and justified their choice of a preferred model for their underlying data, the statistical model preferred by the author team was used, provided the model fulfilled the conditions defined in Section 4.1.5. If several models were appropriate for the underlying data, the simpler model was used, taking into account Section 4.1.5.

4.3.2 Assessment of the informative value of results

The informative value of the results from the included observational studies was assessed on the basis of quality criteria developed especially for studies assessing volume-outcome correlations [28-31]. In terms of the informative value of results, the assessment considered the way the risk adjustment was performed, i.e. the risk factors taken into account and the sources used (administrative databases, clinical databases, medical records). Likewise, the quality of the statistical models used to examine any correlation between VoS and outcome was assessed; this quality depends on the form in which the characteristic of volume entered into the analysis (continuous versus categorical data), the consideration of cluster effects (see Section 4.1.5), and the examination of model quality [32]. The completeness of reporting (e.g. description of analysed data and reporting of point estimates, confidence intervals, and p-values) was considered an aspect of the informative value of results as well. On the basis of the entirety of these quality criteria, the observational studies were categorized by quality into those with high versus low informative value of results.

4.3.3 Assessment of the risk of bias

The risk of bias of the results of the included controlled interventional studies was assessed in accordance with IQWIG General Methods Version 5.0, Chapter 9 [27].

4.3.4 Summary assessment of information

The results on the outcomes reported in the studies were comparatively described in the report.

Since categorical analysis is associated with a loss of information (e.g. the linearity assumption may be violated within the individual categories) and might supply less reliable results than continuous analysis [31], the results of continuous modelling were preferred over those of categorical modelling and included in the report if potential non-linear relationships were adequately taken into account in continuous modelling. However, if the studies presented results exclusively for categorical analysis or if only the results of categorical analysis were usable, the summary assessment used these categorical analyses.

Beyond the comparison of results from the individual studies, suitable metaanalytical methods were to be used if possible [27]. A final summary assessment of the information was performed in any case. Results reported on subgroups were to be presented separately and summarized, if possible.

5 Results

5.1 Comprehensive information retrieval

5.1.1 Primary information sources

Figure 1 shows the results of the systematic literature search in the bibliographic databases and the study selection in accordance with the criteria for study inclusion. The search strategies for the search in bibliographic databases are found in Appendix A. The most recent search was conducted on 20 May 2019.

The references of the hits which were screened at full-text level but excluded are found in Section 9.3 of the full report, with the respective reason for exclusion.

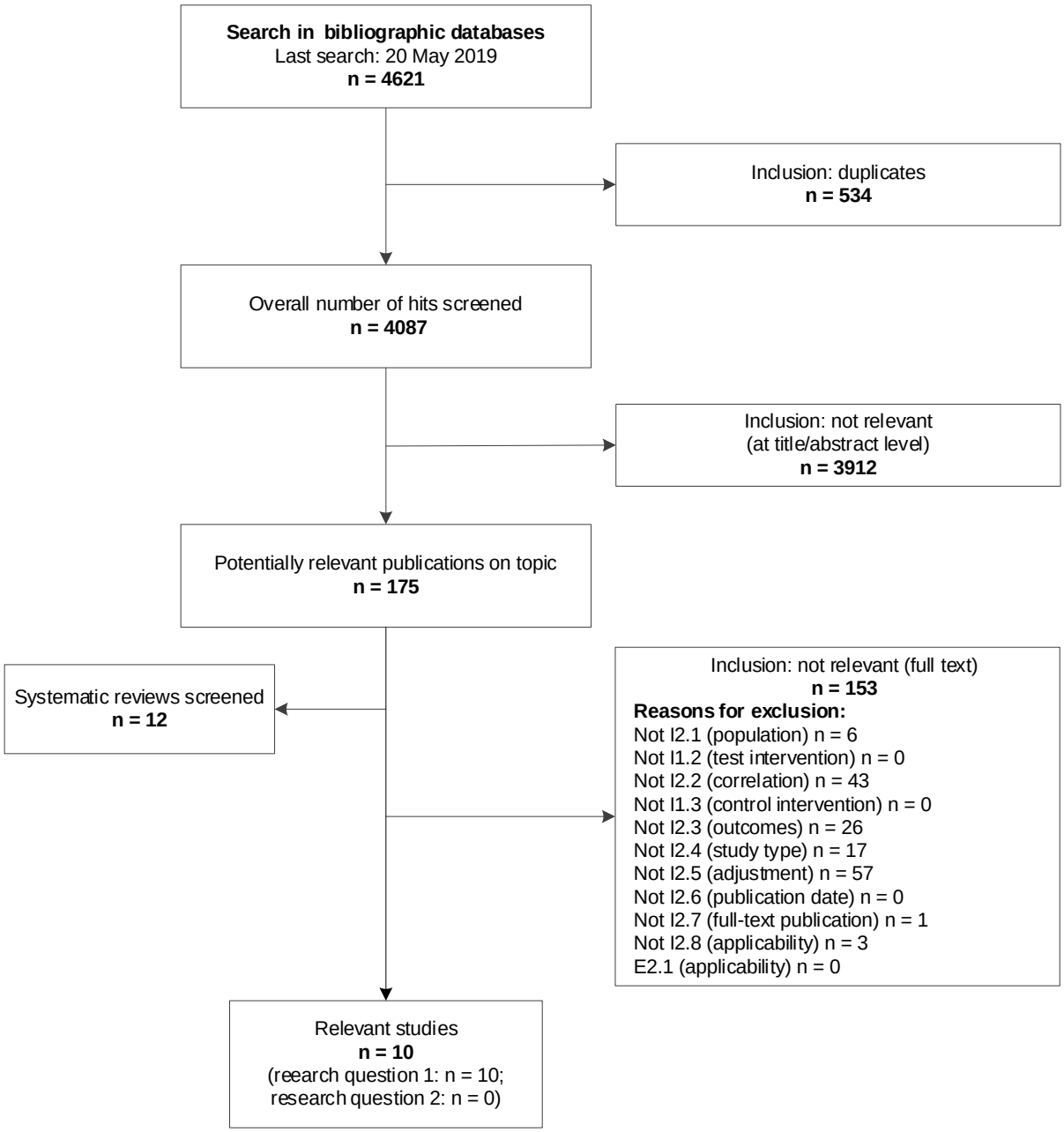

Figure 1: Result of the bibliographic search and study selection

5.1.2 Further information sources and search techniques

Relevant studies or documents found through further information sources and search techniques are presented below unless they were already found through primary information sources.

5.1.2.1 Use of further search techniques

As part of the information retrieval, 12 systematic reviews were found – the corresponding references are provided in Section 9.2 of the full report. The list of references of this systematic review was screened.

No relevant studies or documents which were not already identified in other search steps were found.

5.1.2.2 Requests to authors

No requests to authors to obtain additional information on relevant studies were necessary since such information was not expected to have a relevant impact on the assessment.

5.2 Resulting study pool

Through the various search steps, a total of 10 relevant studies (10 documents) were found (also see Table 3), all of which related to research question 1. The corresponding references are found in Section 9.1 of the full report.

Table 3: Study pool

Study	Full publication (in professional journals)
Allgood 2006	Yes [33]
Gilligan 2007	Yes [34]
Greenup 2018	Yes [35]
Gutierrez 2008	Yes [36]
Isaacs 2016	Yes [37]
McCahill 2012	Yes [38]
Nattinger 2007	Yes [39]
Simunovic 2006	Yes [40]
van Leeuwen 2018	Yes [41]
Vrijens 2012	Yes [42]

No studies of meaningful informative value were found to answer research question 2.

5.3 Characteristics of the studies included in the assessment

The included studies' characteristics which were relevant for the report are presented in Table 4 to Table 8 and then summarized.

Table 4: Characteristics of the included studies

No.	Study / study design (data source)	Study objective	Recruitment country / data collection period / follow-up period	Definition of VoS / threshold definition	Analysis of VoS / number of units in total and, if applicable, per VoS
1	**Allgood 2006** Retrospective observational study (BTW database and NHS Central Register)	Investigation of any correlation between VoS and the care process and/or the outcome of breast cancer patients who were diagnosed through the screening programme	Wales, United Kingdom Screening period: 02/1989–03/1997 Follow-up period [median]: 3.0 years [1.8; 4.2][a]	Number of newly diagnosed and treated women with breast cancer per doctor or per hospital and year Continuous analysis	Analysis: per increase in VoS by 10 patients Hospitals total: 19 Doctors total: 25 Low VoS, ≤ 10 patients: 15 Intermediate VoS, 11–49 patients: 4 High VoS, ≥ 50 patients: 6 Patients total: 2704[b] Low VoS, ≤ 10 patients: 76 (2.8%[c]) Moderate VoS, 11–49 patients: 536 (19.8%[c]) High VoS, ≥ 50 patients: 2092 (77.4%[c])
2	**Gilligan 2007** Retrospective observational study (SEER database and Medicare claims data)	Investigation of any correlation between a hospital's VoS and overall survival or disease-specific survival in breast cancer patients (stages I and II)	USA Period of diagnosis: 1994–1996 Follow-up period [mean]: 62.5 months	Number of surgeries performed annually during the study period in Medicare-insured women with newly diagnosed breast cancer, regardless of stage, per hospital Terciles	Analysis: based on the lowest VoS category Hospitals total: 457 Low VoS, 0–19 patients: 327 Intermediate VoS, 20–39 patients: 87 High VoS, ≥ 40 patients: 43 Patients total: 11 225 Low VoS, 0–19 patients: 3596 (32.0%[c]) Intermediate VoS, 20–39 patients: 3698 (32.9%[c]) High VoS ≥ 40 patients: 3931 (35.0%[c])

(continued)

Table 4: Characteristics of the included studies (continued)

No.	Study / study design (data source)	Study objective	Recruitment country / data collection period / follow-up period	Definition of VoS / threshold definition	Analysis of VoS / number of units in total and, if applicable, per VoS
3	**Greenup 2018** Retrospective observational study (NCDB database)	Investigation of any correlation between a hospital's VoS and overall survival in breast cancer patients and investigation of which VoS threshold value affects survival	USA Period of diagnosis: 2004–2012 Follow-up period: maximum of 11 years	Number of treated breast cancer cases per hospital and year Restricted cubic splines analysis	Analysis: based on the lowest VoS category Hospitals total: 1277[c] Low VoS, < 148 patients: 1044 Intermediate VoS, 148–298 patients: 181 High VoS, > 298 patients: 52 Patients total: 1 064 251 Low VoS, < 148 patients: 546 005 (51.3%) Intermediate VoS, 148–298 patients: 327 405 (30.8%) High VoS, > 298 patients: 190 841 (17.9%)
4	**Gutierrez 2008** Retrospective observational study (FCDS and SEER databases)	Investigation of any correlation between a hospital's VoS or the hospital type and the outcome (survival) of patients with IDC	USA Surgery performance period: 1994–2000 Data collection period: 2006 Follow-up period: 5 and 10 years	Number of surgeries performed in patients with IDC per hospital during the study period 50th percentile (median)	Analysis: based on the lowest VoS category Hospitals total: 296[c] Patients total: 24 834[d] Low VoS, < 50th percentile: 16 147 patients High VoS, > 50th percentile: 7598 patients

(continued)

Table 4: Characteristics of the included studies (continued)

No.	Study / study design (data source)	Study objective	Recruitment country / data collection period / follow-up period	Definition of VoS / threshold definition	Analysis of VoS / number of units in total and, if applicable, per VoS
5	**Isaacs 2016** Retrospective observational study (New York State Mandatory Reporting Database)	Investigation of any correlation between doctor experience and reoperation. In addition, the use of BCS, reoperation rates, and the procedure choice were presented.	USA Surgery performance period: 01/2003–12/2013 Follow-up period: 90 days after surgery	Average number of breast cancer patients treated annually per doctor Terciles	Analysis: based on the highest VoS category Patients total: 89 290[e] Low VoS, 0–13: 30 731 patients (34.4%) Intermediate VoS, 14–33: 28 950 patients (32.4%) High VoS, ≥ 34 patients: 29 609 (33.2%)
6	**McCahill 2012** Retrospective observational study (BRCASO and CRN databases)	Investigation of any correlation between hospital-specific and doctor-specific variation and re-excision following partial mastectomy	USA Period of diagnosis: 2003–2008 N/A	Mean number of performed breast cancer surgeries per year, based on the BRCASO cohort per doctor Quartiles, mean threshold value shifted toward mean	Analysis: based on the highest VoS category Patients total: 1909[f] Low VoS, 0–9.9 patients: 418 Intermediate VoS, 10.0–24.9 patients: 815 High VoS, 25.0–49.9 patients: 178 Very high VoS, ≥ 50.0 patients: 498

(continued)

Table 4: Characteristics of the included studies (continued)

No.	Study / study design (data source)	Study objective	Recruitment country / data collection period / follow-up period	Definition of VoS / threshold definition	Analysis of VoS / number of units in total and, if applicable, per VoS
7	**Nattinger 2007** Retrospective observational study (SEER database and Medicare claims data)	Investigation of any correlation between the doctor's VoS and breast cancer mortality as well as all-cause mortality. In addition, potential bias in the doctor VoS-outcome relationship are analysed.	USA Period of diagnosis 01/1994–12/1996 Median follow-up [median]: 50 months, up to 08/2000	Mean number of breast cancer patients operated annually per doctor or per hospital Definition of thresholds: Categorization into 3 groups, with the highest group defined as having at least double the VoS of doctors of the lowest group	Analysis: based on the lowest VoS category Hospitals total: n.s Low VoS, < 20 patients[h]: 3592[c] (29.4%) Intermediate VoS, 20 to < 40 patients[h]: 3701[c] (30.3%) High VoS, ≥ 40 patients[h]: 3934[c] (32.2%) Unknown: 989[c] (8.1%) Doctors total: 1856 Low VoS, < 5 patients: 1325 (71.4%[c]) Intermediate VoS, 5 to < 10 patients: 384 (20.7%[c]) High VoS, ≥ 10 patients: 147 (7.9%[c]) Patients total: 12 216 Low VoS, < 5 patients[i]: 4524 (37.0%[c]) Intermediate VoS, 5 to < 10 patients[i]: 4456 (36.5%[c]) High VoS, ≥ 10 patients[i]: 3236 (26.5%[c])

(continued)

Table 4: Characteristics of the included studies (continued)

No.	Study / study design (data source)	Study objective	Recruitment country / data collection period / follow-up period	Definition of VoS / threshold definition	Analysis of VoS / number of units in total and, if applicable, per VoS
8	**Simunovic 2006** Retrospective observational study OCR register, CIHI, and ORP databases)	Investigation of any correlation between VoS, hospital teaching status, and intraoperative mortality as well as long-term survival of breast cancer patients.	Canada Period of diagnosis: 01/1991–12/1993 Follow-up period: until 12/2000	Number of cases diagnosed between 1991 and 1993 Definition of thresholds: Quartiles	Analysis: based on the highest VoS category Hospitals total: 152 Patients total: 14 346 Low VoS, ≤ 102 patients: 3569 (24.9%) Low to intermediate VoS, 103–158 patients: 3540 (24.7%) Intermediate to high VoS, 159–264 patients: 3603 (25.1%) High VoS, ≥ 265 patients: 3634 (25.3%)
9	**van Leeuwen 2018** Retrospective observational study (APDC and RBDM registries)	Investigation of any correlation between hospital-specific differences and reoperation within 90 days after breast-conserving surgery	Australia Period of first surgery performance: 07/2002–03/2013 Recruitment period: 07/2001–03/2014 Follow-up period: 90 days after surgery; 12 months before surgery	Average number of breast cancer patients treated annually per hospital. Definition of thresholds: Terciles on the basis of the graphic illustration of distribution	Analysis: based on the lowest VoS category Hospitals total: 161 Patients total: 34 458 Low VoS, < 15 patients: 3278 (9.5%) Intermediate VoS, 15–49 patients: 12 224 (35.5%) High VoS, ≥ 50 patients: 18 956 (55.0%)

(continued)

Table 4: Characteristics of the included studies (continued)

No.	Study / study design (data source)	Study objective	Recruitment country / data collection period / follow-up period	Definition of VoS / threshold definition	Analysis of VoS / number of units in total and, if applicable, per VoS
10	**Vrijens 2012** Retrospective observational study and feasibility study (Belgian cancer registry database, Belgian population database, and administrative claims database)	Investigation of any correlation between hospital VoS and overall survival (5-year survival) in breast cancer as well as 11 quality indicators	Belgium Period of diagnosis: 01/2004–12/2006 Follow-up period: 3 or 5 years (until 12/2009)	Average number of breast cancer patients treated annually per hospital within the study period Quartiles, high VoS according to EUSOMA requirements	Analysis: based on the highest VoS category Hospitals total: 111 Patients total: 25 178 Very low VoS, < 50 patients: 5036 (20.0%) Low VoS, 50–99 patients: 5555 (22.1%) Intermediate VoS, 100–149 patients: 5008 (19.9%) High VoS, ≥ 150 patients: 9579 (38.0%)

a: Interquartile range.
b: Total population: 2705.
c: IQWiG calculation.
d: The analysis used the data of 23 745 patients.
e: The multivariate analysis used the data of at most 63 931 out of 89 448 included patients (due to patient exclusions caused by missing values).
f: Data based only on negative resection margins; total population: 2220 patients.
g: The focus of the analysis is on the doctors' VoS.
h: VoS on the hospital level.
i: VoS on the doctor level.

APDC: NSW Admitted Patient Data Collection; BCS: breast-conserving surgery; BRCASO: Breast Cancer Surgical Outcomes database; BTW: Breast Test Wales; CIHI: Canadian Institute for Health Information; CRN: Cancer Research Network; EUSOMA: European Society of Breast Cancer Specialists; FCDS: Florida Cancer Data System; IDC: Invasive Ductal Carcinoma;; NCDB: National Cancer Data Base; NHS: National Health Service; NSW: New South Wales; OCR: Ontario Cancer Registry; ORPDB: Ontario Registered Persons Database; RBDM: NSW Registry of Births, Deaths and Marriages; SEER: Surveillance Epidemiology and End Results; VoS: volume of services

5.3.1 Study design and data source

A total of 10 studies were included. All studies are retrospective observational studies.

Six of the 10 studies were conducted in the United States (Gilligan 2007, Greenup 2018, Gutierrez 2008, Isaacs 2016, McCahill 2012, Nattinger 2007).

Two U.S. studies (Gilligan 2007, Nattinger 2007) are based on the national data of the Surveillance Epidemiology and End Results Program (SEER) registry, which were linked to administrative data of the U.S. Centers for Medicare and Medicaid Services. The SEER registry contains data on demography, survival, extent of disease, and initial therapy in patients with cancer. Medicare is the U.S. national insurance system which covers old people (65 years and older), people with a disability, and people with dialysis-dependent kidney failure. In 2017, 17.2% of the U.S. population were covered by Medicare [43].

The authors of the Greenup 2018 study, in contrast, used national data from the National Cancer Data Base of the American College of Surgeons. This is a clinical oncology database sourced from hospital registry data that are collected in more than 1500 accredited facilities.

Three studies from the United States included regional data into their analyses. The Gutierrez 2008 study used data from the Florida Cancer Data System. In the Isaacs 2016 study, data from the New York State Mandatory Reporting Data Base were analysed. This administrative database of New York State contains hospital discharge data, data from outpatient surgeries, and data on patient demographics and diagnoses. The McCahill 2012 study is based on data from the Breast Cancer Surgical Outcome Cohort, which was formed by a consortium of the Cancer Research Network and the University of Vermont.

Two further studies were conducted in Europe (Allgood 2006, Vrijens 2012) and are each based on national data.

The Allgood 2006 study used data from the Breast Test Wales (BTW) screening programme and the U.K. National Health Service (NHS). The authors of the Vrijens 2012 study used data based on the Belgian Cancer Registry and the Belgian population database as well as data from a national consortium of all Belgian health insurers. The latter database contains information on all reimbursed drug costs, outpatient and inpatient contacts to doctors, and diagnostic or treatment procedures.

Two further studies are from Canada and Australia (Simunovic 2006, van Leeuwen 2018).

The authors of the Simunovic 2006 study used data from the National Cancer Registry in Ontario, the database of the Canadian Institute for Health Information (CIHI), and the Ontario Registered Persons Database (RPDB). The authors of the van Leeuwen 2018 study examined data from the New South Wales Admitted Patient Data Collection and the New South Wales Registry of Births, Deaths, and Marriages.

5.3.2 Objective of the studies

Seven out of 10 studies examined any correlation between hospital VoS and the mortality of breast cancer patients (Allgood 2006, Gilligan 2007, Greenup 2018, Gutierrez 2008, Nattinger 2007, Simunovic 2006, Vrijens 2012).

Allgood 2006 and Vrijens 2012 also investigated any correlation between VoS and the care process, with Vrijens 2011 examining the care process on the basis of 11 quality indicators, such as the administration of adjuvant systemic therapy. Nattinger 2007 also examined potential influencing factors in the volume-outcome correlation.

Two studies additionally investigated any correlation between VoS and all-cause mortality on the doctor level (Allgood 2006, Nattinger 2007).

In addition to the investigation of any correlation between VoS and mortality, 3 further studies examined the influence of VoS on the morbidity of breast cancer patients (Isaacs 2016, van Leeuwen 2018, McCahill 2012). As their primary objective, the studies investigated any correlation between VoS and reoperation. Two studies based their analyses on the doctor VoS (Isaacs 2016, McCahill 2012), while van Leeuwen 2018 looked at the hospital VoS.

5.3.3 Follow-up period

The follow-up periods of the included studies ranged from 90 days (Isaacs 2016, van Leeuwen 2018) to 11 years after surgery (Gutierrez 2008, Greenup 2018).

The van Leeuwen 2018 and Isaacs 2016 studies are based on data collected between 2002 and 2013. The investigations by Greenup 2018 use data from 2004 through 2012.

Gutierrez et al. investigated data from 1994 through 2000, with a follow-up period of 5 or 10 years. The authors of the Vrijens 2012 study used data from 2004 through 2006, with a maximum follow-up period of 5 years.

Three studies are based on data collected between 1989 and 1997 (Allgood 2006) or 1994 and 1996 (Gilligan 2007, Nattinger 2007). All 3 studies report the median follow-up period, which is between 3 years (Allgood 2006) and 62.5 months (Gilligan 2007). Simunovic et al. included data from 1991 through 1993 and reported a follow-up period of up to 10 years.

5.3.4 Definition of VoS

Two studies calculated the VoS from the breast cancer cases annually treated per doctor/hospital site (Allgood 2006, Greenup 2018).

Six further studies (Isaacs 2016, Gilligan 2007, McCahill 2012, Nattinger 2007, van Leeuwen 2018, Vrijens 2012) calculated the VoS as the mean annual number of treated and/or operated breast cancer patients per doctor or hospital site. The studies' calculations are based on their reported study period or cohort.

The remaining 2 studies (Gutierrez 2008, Simunovic 2006) define the VoS as the total number of surgically treated breast cancer cases within a defined period per doctor/hospital. Gutierrez 2008 refers to the number of operated cases within 7 years, and Simunovic 2006, to the number of diagnosed cases within 3 years.

Allgood 2006 is the only included study to analyse the data continuously and per increase in annual VoS by 10 patients, on both the hospital and doctor level.

The analyses of all other 9 studies were based on the highest and lowest VoS categories. Hence, the studies' authors performed categorical analyses. The Gilligan 2007, Greenup 2018, Gutierrez 2008, van Leeuwen 2018, Simunovic 2006, and Vrijens 2012 studies analysed on the hospital level. The analyses of the Isaacs 2016 and McCahill 2012 studies used the doctor VoS. Nattinger 2007 analysed the results for the VoS per hospital and per doctor.

All studies with categorical analysis, except Gutierrez 2008, used specific threshold values to define the VoS categories. Gutierrez 2018 did not mention a specific threshold value, but instead used the median to define the VoS categories. The indicated VoS thresholds on the hospital level were between <15 (van Leeuwen 2018) and <148 (Greenup 2018) patients for the lowest category and between ≥ 40 (Gilligan 2007, Nattinger 2007) and ≥ 298 (Greenup 2018) cases for the highest.

At the doctor level, the thresholds for the lowest VoS category were between ≤ 5 (Nattinger 2007) and ≤ 13 cases (Isaacs 2016), and for the highest, between ≥ 10 (Nattinger 2007) and ≥ 50 patients (McCahill 2012).

Five studies differentiated the VoS on the basis of 3 VoS categories (Gilligan 2007, Greenup 2018, Isaacs 2016, Nattinger 2007, van Leeuwen 2018). Three studies (McCahill 2012, Simunovic 2006, Vrijens 2012) used 4 VoS categories. Another VoS differentiation used 2 categories on the basis of the median (Gutierrez 2008).

5.3.5 Inclusion and exclusion criteria

The main inclusion and exclusion criteria of the studies are listed in Table 5 and summarized below.

All studies except for Allgood 2006, Gutierrez 2008, van Leeuwen 2018, and Vrijens 2012 defined a minimum age 18 years as an inclusion criterion.

Main exclusion criteria of the studies were male sex, surgery or diagnosis outside the specified study period, or not undergoing breast-conserving surgery as the first surgical treatment.

Table 5: Patient inclusion/exclusion criteria of the studies

Study	Main inclusion criteria	Main exclusion criteria
Allgood 2006	▪ Female sex ▪ New breast cancer diagnosis resulting from screening	▪ Patients with symptoms
Gilligan 2007	▪ Female sex ▪ Age ≥ 66 years ▪ Microscopically confirmed breast cancer diagnosis 1994–1996 ▪ Breast cancer stage I or II ▪ BCS or mastectomy (as initial therapy) ▪ Eligibility for Medicare Part A or B ▪ Not enrolled in a Medicare Health Maintenance Organization	▪ Treatment outside the SEER region ▪ Hospital not identifiable
Greenup 2018	▪ Age between 18 and 90 years ▪ Breast cancer diagnosis between 2004 and 2012 ▪ Stage 0–III ▪ In situ or invasive breast cancer ▪ Unilateral breast cancer ▪ BCS, unilateral, or contralateral mastectomy	▪ Age < 18 years and > 90 years ▪ Diagnosed after 2012 ▪ Stage IV ▪ Bilateral cancer ▪ Unclear laterality ▪ Other surgeries
Gutierrez 2008	▪ Female sex/breast ▪ IDC diagnosis between 1994 and 2000 ▪ Invasive breast cancer ▪ Complete staging data ▪ Surgical cases	▪ In situ tumour
Isaacs 2016	▪ Female sex ▪ Age ≥ 20 years ▪ Breast cancer diagnosis ▪ BCS between 2003 and 2013	▪ BCS was not the 1st surgery ▪ Not a resident of New York State ▪ BCS or mastectomy 1 year before treatment ▪ Prior breast cancer
McCahill 2012	▪ Female sex ▪ Age ≥ 18 years ▪ Breast cancer diagnosis between 2003 and 2008 ▪ Invasive ductal or lobular tumour ▪ The 1st breast cancer surgery (partial mastectomy or open breast biopsy) took place at one of the study sites	▪ Male sex ▪ No surgical treatment ▪ Stage IV ▪ DCIS pathologically confirmed ▪ Clinically suspected inflammatory breast cancer ▪ 1st surgery outside the study sites ▪ Neoadjuvant chemotherapy ▪ Preoperatively diagnosed multifocal breast cancer ▪ Prior radiotherapy of the breast ▪ Mastectomy as 1st surgery ▪ Margin status could not be determined ▪ Patients who were treated by a surgeon with fewer than 10 treated cases in the dataset

(continued)

Table 5: Patient inclusion/exclusion criteria of the studies (continued)

Study	Main inclusion criteria	Main exclusion criteria
Nattinger 2007	▪ Female sex ▪ Age ≥ 66 years ▪ Invasive breast cancer, stage I or II (diagnosed 1994–1996) ▪ Performance of surgery / contact with surgeon within 4 months after diagnosis ▪ Eligibility for Medicare Part A or B ▪ Not enrolled in a Medicare Maintenance Organization	▪ Surgery performed outside the SEER region ▪ No surgeon identification number
Simunovic 2006	▪ All patients with new diagnosis of the described cancer types ▪ Period of diagnosis: 1991–1993 ▪ Performed breast cancer surgery	▪ Prior cancer diagnosis, except in patients < 20 years of age
van Leeuwen 2018	▪ Female sex ▪ Age ≥ 16 years ▪ Breast cancer diagnosis of an invasive or in situ carcinoma ▪ Women who received breast-conserving surgery as the 1st therapy at an NSW (Australia) hospital ▪ Surgery (BCS) performed between 1 Jul 2002 and 31 Mar 2014	▪ 1st BCS before 1 Jul 2002 ▪ 1st BCS after 21 Dec 2013 ▪ Death within 90 days ▪ Not resident of NSW
Vrijens 2012	▪ Female sex ▪ Invasive breast cancer ▪ Hospital treatment between 2004 and 2006	▪ Patients who could not be allocated to any hospital

BCS: breast-conserving surgery; DCIS: ductal carcinoma in situ; IDC: invasive ductal carcinoma, N/A: not available; NSW: New South Wales; SEER: Surveillance Epidemiology and End Results (SEER) Programme

5.3.6 Study population

The population characteristics of the studies included to answer research question 1 are presented in Annex B, Tables 18 through 27, of the full report and summarized below.

The number of patients included in the studies ranged from 2206 (McCahill 2012) to 1 064 251 (Greenup 2018). Included patients were aged between 16 and over 66 years. The analyses of all studies except for Greenup 2018 and Simunovic 2006 examined an exclusively female population.

Furthermore, the 10 studies investigated different forms of breast cancer and focused on different surgical procedures for the treatment of breast cancer. Table 6 and Table 7 provide an overview of the included breast cancer stages and the considered surgical procedures.

Table 6: Overview of the breast cancer stages considered in the studies

Study	Included breast cancer stages (specifications)								
	Invasive tumour	DCIS	Unilateral tumour	Bilateral tumour	Stage 0[a]	Stage I[a]	Stage II[a]	Stage III[a]	Stage IV[a]
Allgood 2006	●	●	●	●	●	●	●	●	●
Gilligan 2007	-	-	-	-	-	●	●	-	-
Greenup 2018	●	●	●	-	●	●	●	●	-
Gutierrez 2008	●[c]	-	-	-	-	-	-	-	-
Isaac 2016[b]	●	●	●	●	●	●	●	●	●
McCahill 2012	●	-	-	-	-	-	-	-	-
Nattinger 2007	●	-	-	-	-	●	●	-	-
Simunovic 2006[b]	●	●	●	●	●	●	●	●	●
van Leeuwen 2018	●	●	-	-	-	-	-	-	-
Vrijens 2012	●	-	-	-	-	-	-	-	-

● Data were reported.
- No data were reported.
a: System according to AJCC.
b: Breast cancer not specified.
c: For ductal tumours only.
AJCC: American Joint Committee on Cancer; DCIS: ductal carcinoma in situ

Table 7: Overview of surgical procedures considered in the studies

Study	Surgical procedure considered (specifications)			
	Mastectomy	BCS	Axillary dissection	SLNB
Allgood 2006	●	●	-	-
Gilligan 2007	●	●	-	-
Greenup 2018	●	●	-	-
Gutierrez 2008[a]	●	●	●	●
Isaacs 2016	-	●	-	-
McCahill 2012	-	●	-	●
Nattinger 2007[a]	●	●	●	●
Simunovic 2006[a]	●	●	●	●
van Leeuwen 2018	-	●	-	-
Vrijens 2012[a]	●	●	●	●

● Data were reported.
- No data were reported.
a: Procedure was not specified.
BCS: breast-conserving surgery; SLNB: sentinel lymph node biopsy

5.3.7 Relevant outcomes

All included studies reported usable data on relevant outcomes. Regarding any correlation between VoS and mortality, results were available on overall survival or all-cause mortality and on (non-)breast cancer mortality. For the outcome category of morbidity, studies reported results on reoperation, including the need for re-excision.

None of the included studies supplied results on the outcomes of adverse effects of therapy, development of local recurrence, disease-free survival, or health-related quality of life.

Table 8 below plainly presents all relevant outcomes of the studies.

Table 8: Matrix of the relevant outcomes with reported results

Study	Outcomes								
	Mortality				Morbidity				
	All-cause mortality / overall survival	Breast cancer mortality	Operative in-hospital death	Other-cause mortality	Reoperation in general[a]	Adverse effects of therapy	Local recurrence	Disease-free survival	Health-related quality of life
Allgood 2006	●	●	-	-	-	-	-	-	-
Gilligan 2007	●	●	-	-	-	-	-	-	-
Greenup 2018	●	-	-	-	-	-	-	-	-
Gutierrez 2008	●	-	-	-	-	-	-	-	-
Isaacs 2016	-	-	-	-	●	-	-	-	-
McCahill 2012	-	-	-	-	●	-	-	-	-
Nattinger 2007	●	●	-	●	-	-	-	-	-
Simunovic 2006	●	-	○	-	-	-	-	-	-
van Leeuwen 2018	-	-	-	-	●	-	-	-	-
Vrijens 2012	●	-	-	-	-	-	-	-	-

● Data were reported and were usable.
○ Data were reported but were not usable for the investigation.
- No data were reported.
a: Includes data on reoperation after BCS in general, mastectomy (as reoperation) and re-excision.
BCS: breast-conserving surgery

5.4 Assessment of the informative value of results

Table 9 presents the informative value of results. For 1 study (Greenup 2018), the informative value of results was rated as high. Key factors behind this rating were an appropriate risk adjustment, a large and representative population, adequate statistical methods, and definition of threshold values.

For all other included studies, the informative value of results was rated as low. This rating was largely due to a lack of consideration of relevant risk factors or low quality and incompleteness of the data used.

Table 10 and Table 11 present an overview of the relevant risk factors taken into account in the studies.

Age was a factor used in the risk adjustment of all studies. Further patient factors considered in almost all studies were patient comorbidities and family history. As relevant tumour factors, the analyses particularly took into account tumour size and lymph node status (see Table 10). Three studies adjusted their analyses for the main treatment method, and 2 other studies additionally adjusted for the administration of radiotherapy, chemotherapy, and hormone therapy. Only 2 studies included risk factors on the doctor level in their investigation. While 4 of the included studies adjusted their investigations for hospital type, 5 studies included the location of the respective hospital as a risk factor into their analyses (see Table 11).

All studies described the method used to take into account cluster effects. Allgood 2006 referred only to the STATA statistics software (1999) and the fact that cluster effects were taken into account in the regression model through a standard option of the programme. All studies reported effect estimators and confidence intervals. Three studies additionally reported p-values. Gilligan et al. and Nattinger et al. also reported on significance. In one study (McCahill 2018), model quality was checked.

Table 9: Informative value of results

Study	High quality of individual data	Adequate patient flow	Volume analysis	Plausible procedure for determining the volume threshold	Suitable model class	Adequate consideration of cluster effects	Adequate risk adjustment	Adequate handling of missing data	Information on a check of model quality	Model validation	Information on point estimate including precision	Adequate reporting of relevant aspects	Other aspects	Informative value of results
Allgood 2006	No	Unclear	Continuous	No	Yes	Yes	Yes	Unclear	No	Unclear	Yes	In part	▪ Primary study goal: VoS and treatment quality ▪ Unclear participation requirement ▪ No evidence of selective reporting	Low
Gilligan 2007	No	Yes	Categorical	Yes	Yes	Yes	No	No	No	Unclear	In part	Yes	▪ Primary study goal: VoS and treatment quality ▪ Participation required ▪ No evidence of selective reporting	Low
Greenup 2018	Yes	Yes	Categorical	Yes	Yes	Yes	Yes	No	No	Unclear	In part	Yes	▪ Primary study goal: VoS and treatment quality ▪ Voluntary participation ▪ No evidence of selective reporting	High
Gutierrez 2008	No	Unclear	Categorical	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Yes	Yes	▪ Primary study goal: VoS and treatment quality ▪ Hospitals are legally required to report all cancer cases to the registry ▪ No evidence of selective reporting	Low

(continued)

Table 9: Informative value of results (continued)

Study	High quality of individual data	Adequate patient flow	Volume analysis	Plausible procedure for determining the volume threshold	Suitable model class	Adequate consideration of cluster effects	Adequate risk adjustment	Adequate handling of missing data	Information on a check of model quality	Model validation	Information on point estimate including precision	Adequate reporting of relevant aspects	Other aspects	Informative value of results
Isaacs 2016	No	Yes	Categorical	Yes	Yes	Yes	No	Yes	No	Unclear	In part		▪ Primary study goal: VoS and treatment quality ▪ Participation required ▪ No evidence of selective reporting	Low
McCahill 2012	No	Yes	Categorical	No	Yes	Yes	Yes	Unclear	Yes	Unclear	Yes	Yes	▪ Primary study goal: VoS and treatment quality ▪ Unclear participation requirement ▪ No evidence of selective reporting	Low
Nattinger 2007	No	Yes	Categorical	Yes	Yes	Yes	No	Unclear	No	Unclear	In part	Yes	▪ Primary study goal: VoS and treatment quality ▪ Participation required ▪ No evidence of selective reporting	Low
Simunovic 2006	Yes	Unclear	Categorical	Yes	Yes	Yes	No	Unclear	No	Unclear	Yes	Yes	▪ Primary study goal: VoS and treatment quality ▪ Hospitals are legally required to report to the registry ▪ No evidence of selective reporting	Low

(continued)

Table 9: Informative value of results (continued)

Study	High quality of individual data	Adequate patient flow	Volume analysis	Plausible procedure for determining the volume threshold	Suitable model class	Adequate consideration of cluster effects	Adequate risk adjustment	Adequate handling of missing data	Information on a check of model quality	Model validation	Information on point estimate including precision	Adequate reporting of relevant aspects	Other aspects	Informative value of results
van Leeuwen 2018	Unclear	Yes	Categorical	No	Yes	Yes	No	Unclear	No	Unclear	In part	Yes	▪ VoS and treatment quality not the primary study objective ▪ Participation required ▪ No evidence of selective reporting	Low
Vrijens 2012	Yes	Yes	Categorical	Unclear	Yes	Yes	No	No	No	Unclear	In part	Yes	▪ Primary study goal ▪ Unclear participation requirement ▪ No evidence of selective reporting	Low

VoS: Volume of services

Table 10: Matrix of risk factors considered in the adjustment (patient level)

Study	Considered risk factors																									
	Patients																									
	Age	Sex	Comorbidities	Ancestry	Education level	Place of residence	Country of birth	Regional per capita income	SES	Income	Aboriginal status	Population density	Propensity group membership	Health insurance	Distance to hospital	Breast cancer history	Year of diagnosis	Year of study inclusion	Stage at diagnosis	Tumour size (T status)	Lymph node involvement (N status)	Metastases (M status)	Invasive vs. in situ (invasiveness)	Grading (aggressiveness)	Hormone receptor status	Bilateral vs. ipsilateral
Allgood 2006	●	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	●	-	-	○	○	-	●	○	-	●
Gilligan 2007	●	-	●	●	●	-	-	●	-	-	-	●	●	-	-	-	-	-	-	●	●	-	-	●	●	-
Greenup 2018	●	●	●	●	●	-	-	-	-	●	-	-	-	●	●	-	-	-	●	●	-	-	-	-	●	-
Gutierrez 2008	●	-	-	●	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	●	●	-	●	●	-	-
Isaacs 2016	●	-	●	●	-	-	-	-	-	-	-	-	-	●	-	●	-	●	-	-	-	-	●	-	-	-
McCahill 2012	●	-	-	●	-	-	-	-	-	-	-	-	-	●	-	-	-	-	-	●	-	-	-	-	-	-
Nattinger 2007	●	-	●	●	-	-	-	●	-	-	-	-	●	-	-	-	-	-	-	●	●	-	-	●	●	-
Simunovic 2006	●	●	●	-	-	●	-	-	●	-	-	-	-	-	-	-	-	-	-	●[a]	●[a]	-	-	-	-	-
van Leeuwen 2018	●	-	●	●	-	-	●	-	●	-	●	-	-	-	-	●	-	-	-	-	-	-	●	-	-	-
Vrijens 2012	●	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	●	●	●	-	●	-	-

● Risk factor taken into account in the adjustment.
○ Taken into account in the adjustment only for invasive tumours.
- No adjustment made for this risk factor.
a: Adjustment for only a subset of the cohort.
SES: socioeconomic status

Table 11: Matrix of risk factors taken into account in the adjustment (procedural, provider, and hospital levels)

Study	Considered risk factors														
	Procedure							Treatment provider				Hospital			
	Main treatment method	Year of surgery	Preoperative diagnostics	Screening	Biopsy methods	Performance of radiotherapy	Performance of chemotherapy	Performance of hormone therapy	Sector (outpatient or inpatient)	VoS	Number of years worked	VoS	Mandate (teaching status, care mandate)	Sociogeographic location	Involvement of multiple facilities
Allgood 2006	●	-	-	-	-	-	○	-	-	●	-	●	-	-	-
Gilligan 2007	-	-	-	-	-	-	-	-	-	-	-	●	-	●	-
Greenup 2018	●	-	-	-	-	●	●	●	-	-	-	-	●	●	●
Gutierrez 2008	●	-	-	-	-	●	●	●	-	-	-	-	●	-	-
Isaacs 2016	-	-	-	-	-	-	-	-	●	-	●	-	-	-	-
McCahill 2012	-	-	●	-	-	-	-	-	-	-	-	-	-	●	-
Nattinger 2007	-	-	-	-	-	-	-	-	-	-	-	●	-	●	-
Simunovic 2006	-	-	-	-	-	-	-	-	-	-	-	-	●	-	-
van Leeuwen 2018	-	●	-	-	-	-	-	-	-	-	-	-	●	●	-
Vrijens 2012	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

● Risk factor taken into account in the adjustment.
○ Taken into account in the adjustment only for invasive tumours.
- No adjustment made for this risk factor.

5.5 Results on relevant outcomes

The results on the outcomes relevant for the report are presented below. As described in Section 5.3.7, the informative value of results is high for one study and low for all other usable studies.

5.5.1 Results on mortality

5.5.1.1 Results on the outcome of all-cause mortality

Seven out of the 10 included studies reported results on the outcome of all-cause mortality (see Table 12). One of the studies with results on the outcome of all-cause mortality (Greenup 2018) was assessed as having a high informative value of results.

Results on the hospital level

In the Greenup 2018 study, a comparison of the highest versus lowest VoS categories showed a statistically significant difference in favour of hospitals with a high VoS for all-cause

mortality, at a follow-up of up to 11 years. The VoS was considered high if more than 298 patients were treated annually per hospital site. Low VoS was defined as fewer than 148 treated patients per year and hospital site. A comparison with the intermediate category (148 to 298 patients) was not performed.

Other supportive results on the hospital level were reported by 5 included studies of a low informative value of results. The studies' follow-up periods were between 3 and 10 years. The VoS categories used differ considerably from the threshold values used by Greenup et al.

A statistically significant difference in favour of hospitals with a high VoS was found in the Gilligan 2007 and Gutierrez 2008 studies. The Vrijens 2012 study also reports a statistically significant difference in favour of the highest VoS category in comparison with the two low VoS categories as regards overall survival after 5 years. In a comparison with the intermediate VoS category (100 to 149 patients), however, no statistically significant difference in favour of the higher VoS category was found.

For overall survival, the Simunovic 2006 study showed a statistically non-significant difference between hospital VoS categories. Simunovic et al. report results for a comparison of the low versus the highest VoS categories.

Similarly, in a continuous analysis per VoS increase by 10 patients or for a subpopulation of 2121 patients with exclusively invasive tumours, Allgood et al. were unable to find a statistically significant difference on the hospital level.

Results on the doctor level

For the VoS per doctor, Allgood 2006 and another study of low informative value of results (Nattinger 2007) additionally report the following:

For a continuous VoS increase by 10 patients per doctor, Allgood et al. report a statistically significant difference in favour of a higher VoS. Furthermore, Allgood et al. present results for a subpopulation of 2121 patients with invasive tumours. At a continuous VoS increase by 10 patients per doctor, a statistically significant difference in favour of higher case numbers was reported for this subpopulation as well.

Over a median follow-up period of 50 months, Nattinger 2007 showed a statistically significant difference in favour of the highest VoS category (≥ 10 cases) per doctor only in comparison with the low VoS category. For the intermediate category (5 to 10 cases), it was not possible to report a significant difference.

Summary

In summary, for all-cause mortality, the studies showed a correlation between VoS and treatment quality in favour of higher VoS. The results show lower all-cause mortality in higher VoS categories: Results on statistically significant differences between hospital VoS in favour of a high VoS were reported by one study with high informative value of results (Greenup

2018). The results of five studies of low informative value of results supported or did not call into question those results, but their VoS categories differed considerably. On the doctor level, 2 studies of a low informative value of results revealed a correlation as well, with their results showing lower all-cause mortality in higher VoS categories.

Table 12: Results – all-cause mortality after breast cancer surgery

Study	Definition of outcome Follow-up	N	VoS specification	Mortality rate, raw n (%)	Adjusted hazard ratio [95% CI]; p-value
Allgood 2006	N/A Follow-up [median]: 3 years	**All tumours (DCIS and invasive)**			
		2705	Per VoS increase by 10 patients per year	203 (7.5[a])	**Doctor level:** **0.90 [0.84; 0.97]; 0.008** **Hospital level:** 1.04 [0.98; 1.12]; 0.213
		Invasive tumours			
		2121	Per VoS increase by 10 patients per year	181 (8.5)	**Doctor level[a]:** 0.91 [0.84; 0.98]; 0.011 **Hospital level[b]:** 1.003 [0.996; 1.010]; 0.38
Gilligan 2007	Time from diagnosis to death or end of study Follow-up [mean]: 62.5 months	11 225	VoS per hospital and year: Low VoS, 0–19 patients Intermediate VoS, 20–39 patients High VoS, ≥ 40 patients	 1036 (28.8) 946 (25.6) 919 (23.4)	 Reference N/A[c] 0.83 [0.75; 0.92]; < 0.05
Greenup 2018	Time from diagnosis to death or last follow-up [maximum] 11 years	1 064 251	VoS per hospital and year: Low VoS, < 148 patients Intermediate VoS, 148–298 patients High VoS, > 298 patients	OS raw[d], n (%): 475 024[e] (87) / 382 204[e] (70) 294 665[e] (90) / 245 554[e] (75) 173 665[e] (91) / 146 948[e] (77)	 Reference N/A 0.89 [0.84; 0.96]; N/A
Gutierrez 2008	Time from diagnosis to death or last follow-up [maximum]: 10 years	23 745[f]	VoS per hospital and 7 years: Low VoS, < median High VoS, > median	 12 433[e] (77)[d] / 10 173[e] (63)[d] 6154[e] (81)[d] / 5243[e] (69)[d]	 Reference 0.90 [0.83; 0.98]; 0.02
Nattinger 2007	N/A Follow-up [median]: 50 months	12 216	VoS per doctor and year: Low VoS, 0–4 patients Intermediate VoS, 5–<10 patients High VoS, ≥ 10 patients	 63.4[g, h] 52.5[g, h] 44.7[g, h]	 Reference 0.94 [0.85; 1.03][i]; N/A 0.86 [0.77; 0.97][i]; < 0.05

(continued)

Table 12: Results – all-cause mortality after breast cancer surgery (continued)

Study	Definition of outcome Follow-up	N	VoS specification[a]	Mortality rate, raw n (%)	Adjusted hazard ratio [95% CI]; p-value
Simunovic 2006	**Long-term survival:** Time from hospital admission to death or last follow-up	979[j, k]	<u>VoS per hospital and 3 years:</u> Low VoS, ≤ 102 patients Low to intermediate VoS, 103–158 patients Intermediate to high VoS,159–264 patients High VoS, ≥ 265 patients	 N/A N/A N/A N/A	 1.3 [0.8; 1.9]; 0.27[l] N/A N/A Reference
Vrijens 2012	Overall survival after 5 years	25 178	<u>VoS per hospital and year:</u> Very low VoS, < 50 patients Low VoS, 50–99 patients Intermediate VoS, 100–149 patients High VoS, ≥ 150 patients	<u>OS raw:</u> 3772[e] (74.9) 4377[e] (78.8) 3996[e] (79.8) 8037[e] (83.9)	 1.26 [1.12; 1.42][l], N/A 1.15 [1.01; 1.30][l]; N/A 1.10 [0.98; 1.24][l]; N/A Reference

a: Values taken from table; conflict with information in body of text.
b: Value taken from body of text; information in table is contradictory and apparently faulty.
c: No calculation possible due to violation of the proportional hazards assumption.
d: After 5 years / 10 years.
e: IQWiG calculation.
f: Total number of patients: 24 834.
g: Deaths per 1000 person-years.
h: 2753 total deaths.
i: Contradictory data on effect measures.
j: Analysis excluding patients with events for in-hospital death.
k: Usable results only for a subset of the cohort with adequate adjustment; a total of 14 346 patients were included.
l: Values > 1 mean an advantage for hospitals with high VoS.

CI: confidence interval; N: number of analysed patients; n: number of patients with an event; N/A: not available; n.s.: not significant; OS: overall survival; VoS: volume of services

5.5.1.2 Results on the outcome of breast cancer mortality

On breast cancer mortality, only 3 studies of low informative value of results delivered usable data (see Table 13).

Results on the hospital level

The Gilligan 2007 and Nattinger 2007 studies each presented results on the basis of hospital VoS. In the multivariate analyses, both studies showed differences in favour of hospitals with a high VoS when comparing all VoS categories with the lowest category. In the Gilligan 2007 study, the determined differences in favour of higher VoS were statistically significant.

Results on the doctor level

On the doctor level, none of the differences reported by Nattinger 2007 and Allgood 2006 were statistically significant.

Summary

In summary, for the outcome of breast cancer mortality, it is possible to derive a correlation between hospital VoS and treatment quality in favour of higher VoS. The correlation was derived on the basis of statistically significant results of one study. Another study presented results that did not numerically contradict the former study. Both studies were assessed as delivering a low informative value of results. The results therefore show lower breast cancer mortality in higher VoS categories. On the doctor level, it was not possible to derive any correlation between VoS and treatment quality from any study for this outcome.

Table 13: Results – breast cancer mortality after breast cancer surgery

Study	Definition of outcome Follow-up	N	VoS specification	Mortality rate, raw n (%)	Adjusted hazard ratio [95% CI]; p-value
Allgood 2006	N/A Follow-up [median]: 3 years	**All tumours (DCIS and invasive)**			
		2705	Continuous analysis per increase of VoS by 10 patients per year	120 (4.4)	**Doctor level:** 0.91 [0.82; 1.00]; n.s. **Hospital level:** N/A
		Invasive tumours only			
		2121	Continuous analysis per increase of VoS by 10 patients per year	113 (5.3)	**Doctor level:** 0.991 [0.981; 1.001]; 0.08 **Hospital level:** N/A
Gilligan 2007	Time from diagnosis to death or end of study Follow-up [mean]: 62.5 months	11 225	<u>VoS per hospital and year:</u>		
			Low VoS, 0–19 patients	263[a] (7.3)	Reference
			Intermediate VoS, 20–39 patients	203[a] (5.5)	0.80 [0.66; 0.97]; N/A
			High VoS, ≥ 40 patients	244[a] (6.2)	0.78 [0.64; 0.96]; N/A
Nattinger 2007	N/A Follow-up [median]: 50 months	12 216[b]	<u>VoS per doctor and year:</u>		
		4524	Low VoS, 0–4 patients	17.4[c, d]	Reference
		4456	Intermediate VoS, 5–9 patients	15.7[c, d]	1.00 [0.84; 1.20][e]; n.s.
		3236	High VoS, ≥ 10 patients	13.0[c, d]	0.94 [0.76; 1.16][e]; n.s.
			<u>VoS per hospital[f] and year:</u>		
		3591[a]	Low VoS, < 20 patients	N/A	Reference
		3701	Intermediate VoS, 20–39 patients	N/A	0.83 [0.68; 1.01][e]; N/A
		3934	High VoS, ≥ 40 patients	N/A	0.84 [0.68; 1.03]e; N/A

a: IQWiG calculations.
b: The final analysis is missing 99 cases.
c: Per 1000 person years.
d: 760 total deaths.
e: Contradictory data on effect measures.
f: The focus of the analysis is on doctor VoS.
CI: confidence interval; N: number of analysed patients; n: number of patients with an event; N/A: not available; n.s.: not significant; VoS: volume of services

5.5.1.3 Results on the outcome of other-cause mortality

A study of low informative value of results (Nattinger 2007) presented results on other-cause mortality (see Table 14).

Results on the hospital level:

No statistically significant differences could be reported for hospital VoS.

Results on the doctor level

On the doctor level, for the comparison of the lowest versus highest VoS category, a statistically significant difference in favour of high VoS was shown.

Summary

In summary, for other-cause mortality, data of low informative value of results can be used to derive a correlation between doctor VoS and treatment quality in favour of higher VoS. The results show lower other-cause mortality in higher VoS categories. In contrast, no correlation can be derived for this outcome between hospital VoS and treatment quality.

Table 14: Results – other-cause mortality after breast cancer surgery

Study	Definition of outcome Follow-up	N	VoS specification	Mortality rate, raw n (%)	Adjusted hazard ratio [95% CI]; p-value
Nattinger 2007	N/A Follow-up [median]: 50 months	12 216[a]	<u>VoS per doctor and year:</u> Low VoS, 0–4 patients Intermediate VoS, 5–9 patients High VoS, ≥ 10 patients <u>VoS per hospital[e] and year:</u> Low VoS, < 20 patients Intermediate VoS, 20–39 patients High VoS, ≥ 40 patients	 46.0[b, c] 36.8[b, c] 31.7[b, c] N/A N/A N/A	 Reference 0.91 [0.81; 1.02][d]; N/A 0.86 [0.75; 0.98][d]; < 0.05 Reference 0.99 [0.88; 1.12][d]; N/A 0.89 [0.78; 1.02][d]; N/A

a: The final analysis is missing 99 cases.
b: Per 1000 person years.
c: 1894 total deaths.
d: Contradictory data on effect measures.
e: The focus of the analysis is on doctor VoS.
CI: confidence interval; N: number of analysed patients; n: number of patients with an event; N/A: not available; VoS: volume of services

5.5.1.4 Results on intraoperative in-hospital deaths

Results on operative in-hospital death are available from the Simunovic 2006 study. However, these results were not usable due to a lack of risk adjustment on the tumour level for this outcome's final analysis model.

5.5.2 Results on morbidity

Three studies of low informative value of results reported data on the outcome category of morbidity and the outcome of performance of reoperation (see Table 15).

Results on the hospital level

Van Leeuwen et al. report usable results on the hospital level. Supplementary data were presented both for all reoperations combined and broken down by mastectomy or re-excision as a second operation.

For reoperation in general, the van Leeuwen 2018 study showed a statistically significant difference only in favour of an intermediate VoS in comparison with the lowest VoS category.

Results on the doctor level

For reoperation, Isaacs et al. and McCahill et al. report results on the doctor level.

In the Isaacs 2016 study, a statistically significant difference in favour of a higher VoS was found when comparing the lowest versus the highest VoS category. In the McCahill 2012 study, reoperation was analysed on the basis of re-excision procedures performed. None of the VoS comparisons revealed any statistically significant differences.

Summary

In summary, for reoperation, a correlation was found between hospital VoS and treatment quality in a study of low informative value of results (van Leeuwen 2018), even for the intermediate VoS category. The results show reduced reoperation rates in higher VoS categories. When comparing the lowest versus the highest VoS category, this correlation apparently remains constant or does not numerically contradict the result, but it is no longer statistically significant.

On the basis of one study (Isaacs 2016) of low informative value of results, a correlation between doctor VoS and treatment quality can be derived in favour of higher VoS for this outcome. The Isaacs 2016 study showed lower reoperation rates with increasing VoS. The results of the McCahill 2012 study did not lend themselves to reliable interpretation but did not call into question the results of the Isaacs 2016 study.

Table 15: Results – reoperation after breast cancer surgery

Study	Definition of outcome Follow-up	N	VoS specification	Reoperation rate, raw n (%)	Adjusted odds ratio [95% CI]; p-value
Isaacs 2016	Reoperation (repeat BCS or mastectomy) 90 days after the first surgical breast cancer treatment (BCS)	89 448[a]			
		All reoperations			
		87 344	VoS per doctor and year:		
		30 235	Low VoS, 0–13 patients	10 655 (35.2)	1.49[a] [1.19; 1.87]; N/A
		28 238	Intermediate VoS, 14–33 patients	8354 (29.6)	1.20[a] [0.93; 1.56]; N/A
		28 871	High VoS, ≥ 34 patients	7931 (27.5)	Reference
McCahill 2012	Reexcision following BCS	2220			
		Reexcision			
		1909[b]	VoS per doctor and year:		
		418	1st quartile: low VoS, 0–9.9 patients	70 (16.8)	0.79 [0.31; 2.02]; N/A[c]
		815	2nd quartile: intermediate VoS, 10–24.9 patients	112 (13.7)	0.81 [0.32; 2.06]; N/A[c]
		178	3rd quartile: high VoS, 25.0–49.9 patients	6 (3.4)	1.31 [0.31; 5.58]; N/A[c]
		498	4th quartile: very high VoS, ≥ 50.0 patients	54 (10.8)	Reference

(continued)

Table 15: Results – reoperation after breast cancer surgery (continued)

Study	Definition of outcome Follow-up	N	VoS specification	Reoperation rate, raw n (%)	Adjusted odds ratio [95% CI]; p-value
van Leeuwen 2018	Reoperation (reexcision and/or mastectomy) 90 days after the first BCS	34 458			
		All reoperation procedures			
			VoS per hospital and year:	10 018	
		3278	Low VoS, < 15 patients	1109 (33.8)	Reference
		12 224	Intermediate VoS, 15-49 patients	3373 (27.6)	0.78 [0.64; 0.96]; N/A
		18 956	High VoS, ≥ 50 patients	5536 (29.2)	0.85 [0.66; 1.09]; N/A
		Of which mastectomy as reoperation			
			VoS per hospital and year:	4872	
		3278	Low VoS, < 15 patients	634 (19.3)	Reference
		12 224	Intermediate VoS, 15–49 patients	1721 (14.1)	0.71 [0.59; 0.85]; N/A
		18 956	High VoS, ≥ 50 patients	2517 (13.3)	0.73 [0.57; 0.93]; N/A
		Of which re-excision			
			VoS per hospital and year:	5146	
		3278	Low VoS, < 15 patients	475 (14.5)	Reference
		12 224	Intermediate VoS, 15–49 patients	1652 (13.5)	0.88 [0.69; 1.12]; N/A
		18 956	High VoS, ≥ 50 patients	3019 (15.9)	0.93 [0.69; 1.24]; N/A

a: Results of the final analysis are based on a cohort of at most 63 931 out of 89 448 patients due to patient exclusion resulting from missing values.
b: Data based only on negative resection margins.
c: Cochran Armitage Trend Test result p = 0.92.
BCS: breast-conserving surgery; CI: confidence interval; N: number of analysed patients; n: number of patients with an event; N/A: not available; VoS: volume of services

5.5.3 Results on adverse effects of therapy

None of the included studies reported data on the outcome of adverse effects of therapy.

5.5.4 Results on local recurrences

None of the included studies reported data on the outcome of local recurrence.

5.5.5 Results on disease-free survival

None of the included studies reported data on the outcome of disease-free survival.

5.5.6 Results on health-related quality of life

None of the included studies reported data on the outcome of health-related quality of life.

5.5.7 Metaanalyses

A metaanalytical summary of results was not generated for any of the reported outcomes. Beyond their varying follow-up periods, the studies considerably differed particularly in the thresholds for the VoS categories as well as the adjusted risk factors.

5.5.8 Subgroup characteristics and other effect modifiers

The authors of the Greenup 2018 study (high informative value of results) conducted subgroup analyses for different age groups, tumour stages, hormone receptor statuses, and hormone therapy use, broken down by hospital VoS.

For the various age groups and patients receiving hormone therapy, no statistically significant evidence of effect modification was available on the basis of an interaction test. For tumour stage and hormone receptor status, no interaction test results were reported. It therefore remains unclear to what extent there is an effect modification relative to the stage and hormone receptor status.

In summary, an effect modification concerning any correlation between VoS and treatment quality was not found for any of the investigated subgroups.

Table 16: Results on subgroup characteristics – all-cause mortality after breast cancer surgery

Study	Definition of outcome	N	Subgroup information	Mortality rate, raw n (%)	Adjusted hazard ratio [95% CI]; p-value
Greenup 2018	Time from diagnosis to death or last follow-up [maximum] 11 years	**Age [years]**			Interaction test: not significant (p > 0.01)
		1 064 251			
			18–40[a, b]	N/A	0.90 [0.81; 1.01]; N/A
			41–55[a, b]	N/A	0.87 [0.79; 0.95]; N/A
			56–69[a, b]	N/A	0.89 [0.82; 0.96]; N/A
			70–90[a, b]	N/A	0.90 [0.84; 0.96]; N/A
		Tumour stage			Interaction test: unclear
		1 064 251			
			Stage 0[b]	N/A	0.79 [0.70; 0.89]; N/A
			Stage I[b]	N/A	0.87 [0.80; 0.94]; N/A
			Stage II[a, b]	N/A	0.92 [0.86; 1.00]; N/A
			Stage III[a, b]	N/A	0.95 [0.88; 1.04]; N/A
		Hormone receptor status			Interaction test: unclear
		969 297			
			ER+/PR+[b]	N/A	0.87 [0.81; 0.94]; N/A
			ER+/PR–[b]	N/A	0.89 [0.80; 0.98]; N/A
			ER–/PR+[b]	N/A	0.96 [0.68; 1.09]; N/A
			ER–/PR–[b]	N/A	0.95 [0.89; 1.02]; N/A
		Hormone therapy			Interaction test not significant (p > 0.01)
		1 064 251			
			Yes[a, b]	N/A	0.87 [0.80; 0.94]; N/A
			No[a, b]	N/A	0.90 [0.86; 0.98]; N/A

a: Values read off diagram.
b: High versus low VoS.

CI: confidence interval; ER: oestrogen receptor; N: number of analysed patients; n: number of patients with an event; N/A: not available; PR: progesterone receptor; VoS: volume of services

5.6 Overall evaluation of results

For the outcome of all-cause mortality, a study of high informative value of results (Greenup 2018) revealed statistically significant differences in favour of a higher hospital VoS. Five additional studies of low informative value of results support these results.

Across studies, on the basis of 1 study of high informative value of results and 5 studies of low informative value of results, with studies showing similar results, but having a wide variation of VoS categories, a correlation between hospital VoS and treatment quality can be derived for this outcome. On the doctor level, a correlation was found for all-cause mortality in 2 other studies of low informative value of results. The results show all-cause mortality lower at a higher VoS, on both the hospital and doctor levels.

For other mortality outcomes, results were available from studies of low informative value of results.

For the outcome of breast cancer mortality, a correlation between hospital VoS and treatment quality was also derived on the basis of 2 studies with a follow-up period of 50 to 62.5 months. Across studies, all differences were in favour of a higher hospital VoS, and some were statistically significant. The results therefore show lower breast cancer mortality in higher VoS categories. Across studies, it was not possible to derive any correlation between VoS and treatment quality on the doctor level.

For other-cause mortality, it was possible to derive a correlation between doctor VoS and treatment quality. The latter correlation was derived on the basis of only one study. No correlation between VoS and treatment quality in favour of a higher VoS on the hospital level can be derived for this outcome. The results on the doctor level show a decrease in other-cause mortality.

In terms of morbidity, 3 studies reported the performance of reoperation as an outcome. For the outcome of reoperation in general, a correlation both on the hospital level and on the doctor level can be derived, each from 1 and 2 respective studies of low informative value of results. The results show lower reoperation rates in higher VoS categories. At the hospital level, any correlation between VoS and treatment quality was already derived on the basis of the comparison with the intermediate VoS.

No data were available for further relevant morbidity-related outcomes (e.g. disease-free survival).

No data were reported on health-related quality of life; therefore, for this outcome, no conclusion can be drawn on any correlation between hospital VoS and the quality of treatment outcome.

Since no studies with meaningful interpretive value were found, it was not possible to draw a conclusion on the effects of minimum case numbers introduced for the surgical treatment of breast cancer on the quality of treatment outcomes. Table 17 presents an overview of all results on the outcomes.

Table 17: Overview – correlation between VoS and outcomes

	Outcomes								
Mammary carcinoma surgery	**Mortality**				**Morbidity**				**Health-related quality of life**
	All-cause mortality	**Breast cancer mortality**	**Operative in-hospital death**	**Other-cause mortality**	**Performance of reoperation in general**	**Adverse effects of therapy**	**Local recurrence**	**Disease-free survival**	
Hospital level	↑	(↑)	-	(↔)	(↑)[a]	-	-	-	-
Doctor level	(↑)	(↔)	-	(↑)[a]	(↑)[b]	-	-	-	-
Correlation between VoS and quality of treatment outcome	Correlation between VoS and quality of treatment outcome can be derived on both levels.	Correlation between VoS and quality of treatment outcome can be derived on the hospital level.	No conclusion can be drawn.	Correlation between VoS and quality of treatment outcome can be derived on the doctor level.	Correlation between VoS and quality of treatment outcome can be derived on both levels.	No conclusion can be drawn.			

↑: In 1 study of high informative value of results, a statistically significant difference was found for the outcome in favour of hospitals with higher VoS. Studies of low informative value of results show the same trend or do not call the correlation into question.
(↑): One or more studies of a low informative value of results show, for at least one VoS comparison, a statistically significant difference regarding the outcome in favour of hospitals and/or doctors with high VoS. The numeric advantage in favour of a higher VoS was found across all studies.
(↔): In studies of low informative value of results, no statistically significant differences between the hospitals or doctors with lower versus higher VoS were found for the outcome.
-: The included studies did not report any (usable) results on this outcome.
a: The results are each based on a study of low informative value of results.
b: The results are based on 2 studies of low informative value of results, of which only 1 study permits a conclusive interpretation.
VoS: volume of services

6 Discussion

This report aimed to present and assess a potential correlation between VoS and quality of treatment outcomes for the surgical treatment of breast cancer as well as the effects of specific minimum volumes introduced in the healthcare system on the quality of the treatment outcome. The G-BA commissioned this report against the backdrop of a consultation procedure which was initiated to establish a minimum volume requirement for the surgical treatment of breast cancer.

Usable data were found exclusively on a correlation between VoS and treatment quality. Since no studies of meaningful interpretive value were found, it was not possible to draw a conclusion on the effects of minimum case numbers introduced for the surgical treatment of breast cancer on the quality of treatment outcomes. Likewise, the included studies did not provide any information on palliative surgery cases.

For assessing a correlation, results were available both per hospital and per doctor. The identified studies do not show the extent to which the observed differences in treatment quality were influenced by factors such as internal hospital processes and structures or individual doctor qualifications related to the number of treated breast cancer patients.

Nattinger 2007 states that the correlation observed between VoS and treatment quality for the outcome of mortality is influenced particularly by factors such as patient comorbidities (e.g. cardiovascular death). For example, a significant correlation was found between VoS and treatment quality for the outcome of other-cause mortality, but not for the outcome of breast cancer mortality. The authors conclude that the observed benefit of a higher VoS might be overestimated. In the Greenup 2018 study with a high informative value of results, relevant confounders such as patient comorbidities were taken into account in a multivariate analysis. Furthermore, subgroup analyses of the Greenup 2018 study showed no evidence of an effect modification in any of the considered characteristics (see Section 5.5.8). Therefore, the results of the Greenup 2018 study, which show a correlation between VoS and all-cause mortality, are not considered to be called into question by the results of the Nattinger 2007 study.

It was also found that the influence of VoS is not seen as a factor of a simple causal relationship, but rather part of a multifactorial constellation, depending on which level is considered and which conditions prevail in terms of the process and structure [42, 44].

For instance, Vrijens et al. have shown that treatment quality must be considered in a multifactorial constellation, not least due to the multidisciplinary process. In addition to the correlation between hospital VoS and all-cause mortality, the authors also investigated 11 process indicators. The influence of the VoS on various care processes such as the performance of neoadjuvant therapy or the percentage of breast-conserving operations was examined. Regarding the implementation of evidence-based care processes, a statistically significant difference was found in favour of facilities with a higher VoS for 6 out of 11 indicators [42]. Since no cluster effects were taken into account in the analysis of the implementation of care processes, it was not possible to take them into account in the presentation of results.

Any influence of VoS on treatment quality must be determined while controlling for other influencing factors [31, 44]. Therefore, studies were included only if they included a risk adjustment for relevant confounders. In addition to taking into account established risk factors such as age and sex, it was possible to identify disease-specific factors such as tumour size, tumour stage, or hormone receptor status. Inclusion required that, in addition to general risk factors, at least one disease-specific factor was taken into account in order to reflect the severity of the disease in the analysis. This means that in the study selection, studies which failed to reflect the severity of disease in their adjustment were excluded [45-50]. Since in the Simunovic 2006 study, a risk adjustment by tumour factors was performed only in a subset of the cohort and presented only for the analysis of the outcome of all-cause mortality, the results on the outcome of operative in-hospital death were not usable.

Beyond study inclusion, risk adjustment played a role in the assessment of the informative value of results. The assessment included the extent to which the clinical picture was reflected in the risk adjustment and the levels which were included in the risk adjustment. In this regard, 4 levels (patient characteristics including tumour characteristics, procedure, hospital, and operating doctor) were identified. The informative value of results was rated high if an adequate adjustment existed on the levels relevant for the research question. Accordingly, to be rated as having a high informative value of results, it was sufficient if the levels taken into account included the patient, tumour, procedure, as well as either the hospital or doctor. The Vrijens 2012 study, for instance, was rated as having a low informative value of results since adjustments had been made exclusively on the patient and tumour levels.

In addition to adjusting for relevant risk factors, included studies had to consider cluster effects in their analyses. Some of the publications screened in full text conducted no analysis in consideration of cluster effects on the level of the hospital or the treating doctor and had to be excluded (e.g. [51, 52]). If an analysis fails to take into account cluster effects, patient data are analysed without considering the associated conditions on the hospital or doctor level. This can result in differences being overestimated and confidence intervals being estimated as too narrow since the results of individuals within the same hospital cannot be considered statistically independent [31]. For a study to be included, the consideration of cluster effects on either the hospital or doctor level sufficed. Since the Gentil 2012 study [53] considered cluster effects only in one analysis which was irrelevant for the research question of this report due to the outcomes examined, this study was excluded in the study selection process.

Other studies (e.g. [54-57]) analysing the use of certain treatment procedures or combinations thereof (e.g. BCS and subsequent radiotherapy) as the outcome were excluded. This was because the selection and use of a procedure depend on many factors unrelated to the VoS, and these factors failed to be considered, even after adequate risk adjustment. Such factors include, for instance, hospital structures and the associated ability to offer a treatment or preferences of a patient who is involved in decision-making [19]. It is often impossible to adequately represent the interplay of such factors in a study. When investigating any correlation between VoS and

treatment quality, the use of a specific procedure could therefore not be considered a comparable outcome across studies.

The breakdown into separate procedures was not taken into account in the interpretation of results for the outcome category of morbidity either. In the presentation of results on the outcome of reoperation, the results for different procedures are in fact listed, but an interpretation and the derivation of a correlation are provided only for reoperation in general, regardless whether mastectomy or re-excision were performed. The procedure chosen as the second operation depends not only on the result of the first operation, but also, among other things, on the procedure used for the first surgery and the patient's preference of mastectomy over breast-conserving surgery [19, 58].

In contrast, a quality indicator established in Germany is lymph node dissection in DCIS and breast-conserving therapy [59]. The goal is to achieve the lowest possible percentage of breast cancer patients with DCIS and breast-conserving therapy who underwent axillary lymph node dissection. Among others, 3 studies [46, 47, 60] were identified which investigated any correlation between VoS and the performance of axillary dissection in patients with DCS [46, 47, 60]. These studies had to be excluded due to a lack of risk adjustment for relevant confounders, not least because this quality indicator applies only to certain patients (patients with DCIS and breast-conserving therapy), which has to be taken into account in the selection of the population and the adjustment. Further quality indicators evaluated by IQTIG include performance of primary axillary dissection in DCIS, intraoperative specimen radiography, and determination of the hormone receptor status [59]. Since none of the included studies investigated the quality indicators established in Germany, and the results for the corresponding analysis by Vrijens et al. were not usable, it was not possible to present related results and take them into account when deriving any correlation between VoS and treatment quality [59].

In Germany, minimum case numbers per hospital site (100 primary cases) are not legally mandated but are in place as part of the DKG's certification requirements for breast centres [9, 10]. The DKG requirements are based, among other things, on the quality indicators established by the working group of the European Society of Breast Cancer Specialists (EUSOMA) [61, 62]. It specifies the annual case volume as only one of a variety of requirements a hospital must meet to be certified as a breast centre. Other requirements relate, for instance, to employees' individual qualifications, maintenance of tumour documentation, and the same-site availability of palliative care. Interdisciplinary collaboration is another part of the certification requirements, and various specialties like gynaecology, radiology, or oncology and their cooperation within the care process are taken into account. Accordingly, numerous different aspects of the care process are addressed in the breast cancer certification process with the goal of establishing good treatment quality [9, 10].

The fact that breast centre certification can play a role in the care process and may affect treatment quality has been shown, for instance, by the Schrodi 2015 study [63], which was excluded in this report for its failure to consider cluster effects. Furthermore, Geraedts et al.

[64] investigated changes in the care process due to the establishment of breast centres. The authors of the Geraedts 2013 study reported a lower percentage of breast cancer patients being treated in non-certified centres.

Legally required minimum volumes have been established in France, for instance, where hospitals have to treat at least 30 cases annually [65]. The Rococo 2016 study [66], which was excluded due to its failure to consider cluster effects, investigated to what extent the rates of use of various surgical treatment procedures changed after the introduction of minimum volume rules. The study's authors categorized treating hospitals by size and profile. The study reports that after the introduction of minimum case numbers, larger hospitals more often performed immediate breast reconstructions and sentinel lymph node biopsies than did smaller facilities.

All included studies, except for Allgood 2006, conducted categorical analysis on the basis of VoS categories, which were defined by thresholds. The VoS category thresholds selected by the studies varied greatly, not least because they were formed arbitrarily or using different procedures. Thresholds on the hospital level were between <15 (van Leeuwen 2018) and <148 patients (Greenup 2018) for the lowest category and between ≥ 40 (Gilligan 2007, Nattinger 2007) and ≥ 298 cases (Greenup 2018) for the highest. The thresholds for the lowest VoS per doctor were between ≤ 5 (Nattinger 2007) and ≤ 13 patients (Isaacs 2016), and for the highest category, between ≥ 10 (Nattinger 2007) and ≥ 50 (McCahill 2012).

Further, 8 out of the 10 included studies were conducted outside Europe, 6 of which in the United States [34-41]. Healthcare structures in other countries differ from those in Germany. For instance, the specialties' involvement in the breast cancer treatment process differ between Germany and the USA. Therefore, the studies are not fully transferable to the German healthcare system. Subgroup analyses in the Greenup 2018 study (high informative value of results) showed that characteristics such as differences in the healthcare structure (health insurance) do not represent effect modifiers. Therefore, the results do not necessarily entirely rule out transferability to the German healthcare context.

Likewise, some of the included studies reported results on data which were more than 20 years old when the report was generated. Hence, the question is to what extent their results adequately reflect current treatment and quality standards and advances in breast cancer treatment.

In summary, on the basis of the included studies, it was possible to derive a correlation between VoS and treatment quality, particularly for all-cause mortality since a study of high informative value of results was available on this outcome. These study results show lower all-cause mortality in higher VoS categories. For other outcomes, correlations were derived as well, but they are based on studies of a low informative value of results. Again, the results showed lower mortality or morbidity, either on the hospital or doctor level, in higher VoS categories. Studies with meaningful interpretive value would be helpful, particularly in case of the introduction of minimum volume rules, in order to be able to determine any influence of the VoS and specific effects of any minimum volume introduced.

7 Conclusion

For the investigation of any correlation between VoS and treatment quality in the surgical treatment of breast cancer, a total of 10 studies were included. One of the 10 included studies was assessed as having a high informative value of results.

For the outcome of all-cause mortality, based on a high informative value of results, a correlation between both hospital and doctor VoS and treatment quality was derived. Hence, a higher VoS is associated with lower mortality rates. For other mortality outcomes (breast cancer mortality and other-cause mortality), based on a low informative value of results, a correlation was found either only on the hospital level or only on the doctor level. Again, the results showed lower mortality for these levels.

With respect to performing a reoperation, based on a low informative value of results, it was possible to derive a correlation between both doctor and hospital VoS and treatment quality. At the hospital level, a correlation between VoS and treatment quality was apparent even in comparison with the intermediate VoS category. The results show lower reoperation rates in higher VoS categories.

No studies of meaningful interpretive value were found which investigated any correlation between VoS and treatment quality reflected by other outcomes, such as adverse effects of therapy, local recurrence, disease-free survival, or health-related quality of life. Further, it was not possible to include studies of meaningful interpretive value investigating any effects of specifically introduced minimum case numbers.

References

1. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? N Engl J Med 1979; 301(25): 1364-1369.

2. Chowdhury MM, Dagash H, Pierro A. A systematic review of the impact of volume of surgery and specialization on patient outcome. Br J Surg 2007; 94(2): 145-161.

3. Loberiza FR Jr, Zhang MJ, Lee SJ, Klein JP, LeMaistre CF, Serna DS et al. Association of transplant center and physician factors on mortality after hematopoietic stem cell transplantation in the United States. Blood 2005; 105(7): 2979-2987.

4. Gandjour A, Bannenberg A, Lauterbach KW. Threshold volumes associated with higher survival in health care: a systematic review. Med Care 2003; 41(10): 1129-1141.

5. Killeen SD, O'Sullivan MJ, Coffey JC, Kirwan WO, Redmond HP. Provider volume and outcomes for oncological procedures. Br J Surg 2005; 92(4): 389-402.

6. Matthias K, Gruber S, Pietsch B. Evidenz von Volume-Outcome-Beziehungen und Mindestmengen: Diskussion in der aktuellen Literatur. Gesundheits- und Sozialpolitik 2014; (3): 23-30.

7. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [online]. URL: https://www.g-ba.de/informationen/richtlinien/42/.

8. Gemeinsamer Bundesausschuss. Mindestmengenregelungen gemäß § 136b Abs. 1 Satz 1 Nr. 2 SGB V [online]. URL: https://www.g-ba.de/richtlinien/5/.

9. Mansky T, Drogan D, Nimptsch U, Günster C. Stationäre Versorgungsstruktur ausgewählter Krankeheitsbilder in Deutschland. In: Dormann F, Klauber J (Ed). Qualitätsmonitor 2017. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2017. p. 163-320. URL: https://www.wido.de/fileadmin/Dateien/Dokumente/Publikationen_Produkte/Buchreihen/Qualitaetsmonitor/wido_ver_qualitaetsmonitor_2017_gesamt_1116.pdf.

10. Deutsche Krebsgesellschaft. Jahresbericht 2018 der zertifizierten onkologischen Zentren. Berlin DKG; 2018.

11. Zentrum für Krebsregisterdaten. Brustkrebs (Mammakarzinom): ICD-10 C50 [online]. 06.12.2017 [Accessed: 25.03.2019]. URL: https://www.krebsdaten.de/Krebs/DE/Content/Krebsarten/Brustkrebs/brustkrebs.html.

12. Wörmann B, Aebi S, Balic M, Decker T, Fehm T, Greil R et al. Mammakarzinom des Mannes [online]. In: Onkopedia Leitlinien. 08.2016 [Accessed: 26.11.2019]. URL: https://www.onkopedia.com/de/onkopedia/guidelines/mammakarzinom-des-mannes/@@pdf-latest?filename=mammakarzinom-des-mannes.pdf.

13. Eucker J, Regierer AC, Possinger K. Mammakarzinome. In: Possinger K, Regierer AC, Eucker J (Ed). Facharztwissen Hämatologie Onkologie. München: Elsevier; 2017. p. 787-825.

14. Janni W, Müller V (Ed). Brustkrebs: Patientenratgeber zu den AGO Empfehlungen 2019. München: Zuckschwerdt; 2019. URL: https://www.ago-online.de/fileadmin/downloads/pdf/2019/AGO_Brustkrebs_2019.pdf.

15. Sun YS, Zhao Z, Yang ZN, Xu F, Lu HJ, Zhu ZY et al. Risk factors and preventions of breast cancer. Int J Biol Sci 2017; 13(11): 1387-1397.

16. Zentrum für Krebsregisterdaten, Gesellschaft der Epidemiologischen Krebsregister in Deutschland. Krebs in Deutschland für 2013/2014. Berlin: RKI; 2017. URL: https://www.krebsdaten.de/Krebs/DE/Content/Publikationen/Krebs_in_Deutschland/kid_2017/krebs_in_deutschland_2017.pdf?__blob=publicationFile.

17. Deutsche Krebsgesellschaft. Brustkrebs Diagnose: Klassifikation [online]. 21.02.2017 [Accessed: 26.04.2019]. URL: https://www.krebsgesellschaft.de/onko-internetportal/basis-informationen-krebs/krebsarten/brustkrebs/klassifikation.html.

18. Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ et al. Breast cancer: major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin 2017; 67(4): 290-303.

19. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften, Deutsche Krebsgesellschaft, Deutsche Krebshilfe. Interdisziplinäre S3-Leitlinie für die Früherkennung, Diagnostik, Therapie und Nachsorge des Mammakarzinoms; Langversion 4.2; AWMF-Registernummer 032-045OL [online]. 08.2019 [Accessed: 26.11.2019]. URL: https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/Downloads/Leitlinien/Mammakarzinom_4_0/Version_4.2/LL_Mammakarzinom_Langversion_4.2.pdf.

20. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. N Engl J Med 2002; 347(16): 1233-1241.

21. Veronesi U, Cascinelli N, Mariani L, Greco M, Saccozzi R, Luini A et al. Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. N Engl J Med 2002; 347(16): 1227-1232.

22. Deutsche Krebsgesellschaft. Brustkrebs: Operationen [online]. 23.02.2017 [Accessed: 26.04.2019]. URL: https://www.krebsgesellschaft.de/onko-internetportal/basis-informationen-krebs/krebsarten/brustkrebs/therapie/operation.html.

23. Wörmann B, Aebi S, Balic M, Decker T, Fehm T, Greil R et al. Mammakarzinom der Frau [online]. In: Onkopedia Leitlinien. 01.2018 [Accessed: 26.11.2019]. URL: https://www.onkopedia.com/de/onkopedia/guidelines/mammakarzinom-der-frau/@@pdf-latest?filename=mammakarzinom-der-frau.pdf.

24. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3; current step 4 version [online]. 30.11.1995 [Accessed: 05.09.2018]. URL: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf.

25. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health 2004; 94(3): 361-366.

26. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577.

27. 1. Institute for Quality and Efficiency in Health Care. General methods: version 5.0 [online]. 10.07.2017 [Accessed: 13.09.2018]. URL: https://www.iqwig.de/download/General-Methods_Version-5-0.pdf.

28. Bender R, Grouven U. Possibilities and limitations of statistical regression models for the calculation of threshold values for minimum provider volumes. Z Evid Fortbild Qual Gesundhwes 2006; 100(2): 93-98.

29. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Entwicklung und Anwendung von Modellen zur Berechnung von Schwellenwerten bei Mindestmengen für die Koronarchirurgie: Vorbericht; Auftrag B05/01b [online]. 24.04.2006 [Accessed: 11.03.2013]. URL: http://www.iqwig.de/download/B05-01B_Vorbericht_Entwicklung_und_Anwendung_von_Modellen_zur_Berechnung_von_Schwellenwerten_bei_Mindestmengen_fuer_die_Koronarchirurgie.pdf.

30. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Zusammenhang zwischen Menge der erbrachten Leistungen und der Ergebnisqualität für die „Perkutane Transluminale Coronare Angioplastie (PTCA)“: Abschlussbericht; Auftrag Q05-01B [online]. 06.06.2006 [Accessed: 11.03.2013]. (IQWiG-Berichte; Volume 8). URL: https://www.iqwig.de/download/Q05-01B_Abschlussbericht_Zusammenhang_Menge_erbrachter_Leistung_und_Ergebnisqualitaet_bei_PTCA..pdf.

31. Wetzel H. Definition of low threshold volumes for quality assurance: conceptual and methodological issues involved in the definition and evaluation of thresholds for volume outcome relations in clinical care. Z Arztl Fortbild Qualitatssich 2006; 100(2): 99-106.

32. Hosmer DW, Lemeshow S. Applied logistic regression. New York: Wiley; 2000.

33. Allgood PC, Bachmann MO. Effects of specialisation on treatment and outcomes in screen-detected breast cancers in Wales: cohort study. Br J Cancer 2006; 94(1): 36-42.

34. Gilligan MA, Neuner J, Zhang X, Sparapani R, Laud PW, Nattinger AB. Relationship between number of breast cancer operations performed and 5-year survival after treatment for early-stage breast cancer. Am J Public Health 2007; 97(3): 539-544.

35. Greenup RA, Obeng-Gyasi S, Thomas S, Houck K, Lane WO, Blitzblau RC et al. The effect of hospital volume on breast cancer mortality. Ann Surg 2018; 267(2): 375-381.

36. Gutierrez JC, Hurley JD, Housri N, Perez EA, Byrne MM, Koniaris LG. Are many community hospitals undertreating breast cancer? Lessons from 24,834 patients. Ann Surg 2008; 248(2): 154-162.

37. Isaacs AJ, Gemignani ML, Pusic A, Sedrakyan A. Association of breast conservation surgery for cancer with 90-day reoperation rates in New York State. JAMA Surg 2016; 151(7): 648-655.

38. McCahill LE, Single RM, Aiello Bowles EJ, Feigelson HS, James TA, Barney T et al. Variability in reexcision following breast conservation surgery. JAMA 2012; 307(5): 467-475.

39. Nattinger AB, Laud PW, Sparapani RA, Zhang X, Neuner JM, Gilligan MA. Exploring the surgeon volume-outcome relationship among women with breast cancer. Arch Intern Med 2007; 167(18): 1958-1963.

40. Simunovic M, Rempel E, Theriault ME, Coates A, Whelan T, Holowaty E et al. Influence of hospital characteristics on operative death and survival of patients after major cancer surgery in Ontario. Can J Surg 2006; 49(4): 251-258.

41. Van Leeuwen MT, Falster MO, Vajdic CM, Crowe PJ, Lujic S, Klaes E et al. Reoperation after breast-conserving surgery for cancer in Australia: statewide cohort study of linked hospital data. BMJ Open 2018; 8(4): e020858.

42. Vrijens F, Stordeur S, Beirens K, Devriese S, Van Eycken E, Vlayen J. Effect of hospital volume on processes of care and 5-year survival after breast cancer: a population-based study on 25000 women. Breast 2012; 21(3): 261-266.

43. Berchick ER, Hood E, Barnett JC. Health insurance coverage in the United States: 2017 [online]. 12.09.2018 [Accessed: 17.10.2019]. (Current Population Reports; Volume P60-264). URL: https://www.census.gov/content/dam/Census/library/publications/2018/demo/p60-264.pdf.

44. Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. Planungsrelevante Qualitätsindikatoren: Prüfung der Ableitung aus Richtlinien zur Strukturqualität und Mindestmengenregelungen; Abschlussbericht [online]. 27.04.2018 [Accessed: 01.03.2019]. URL: https://iqtig.org/dateien/berichte/2018/IQTIG_PlanQI-Folgeauftrag-1_Abschlussbericht_2018-04-27.pdf.

45. Avdic D, Lundborg P, Vikström J. Estimating returns to hospital volume: evidence from advanced cancer surgery. J Health Econ 2019; 63: 81-99.

46. Coromilas EJ, Wright JD, Huang Y, Feldman S, Neugut AI, Chen L et al. The influence of hospital and surgeon factors on the prevalence of axillary lymph node evaluation in ductal carcinoma in situ. JAMA Oncol 2015; 1(3): 323-332.

47. Gilligan MA, Neuner J, Sparapani R, Laud PW, Nattinger AB. Surgeon characteristics and variations in treatment for early-stage breast cancer. Arch Surg 2007; 142(1): 17-22.

48. Hynes DM, Weaver F, Morrow M, Folk F, Winchester DJ, Mallard M et al. Breast cancer surgery trends and outcomes: results from a National Department of Veterans Affairs study. J Am Coll Surg 2004; 198(5): 707-716.

49. Yen TWF, Pezzin LE, Li J, Sparapani R, Laud PW, Nattinger AB. Effect of hospital volume on processes of breast cancer care: a national cancer data base study. Cancer 2017; 123(6): 957-966.

50. Yen TWF, Li J, Sparapani RA, Laud PW, Nattinger AB. The interplay between hospital and surgeon factors and the use of sentinel lymph node biopsy for breast cancer. Medicine (Baltimore) 2016; 95(31): e4392.

51. Peltoniemi P, Peltola M, Hakulinen T, Häkkinen U, Pylkkänen L, Holli K. The effect of hospital volume on the outcome of breast cancer surgery. Ann Surg Oncol 2011; 18(6): 1684-1690.

52. Siesling S, Tjan-Heijnen VCG, De Roos M, Snel Y, Van Dalen T, Wouters MW et al. Impact of hospital volume on breast cancer outcome: a population-based study in the Netherlands. Breast Cancer Res Treat 2014; 147(1): 177-184.

53. Gentil J, Dabakuyo TS, Ouedraogo S, Poillot ML, Dejardin O, Arveux P. For patients with breast cancer, geographic and social disparities are independent determinants of access to specialized surgeons: a eleven-year population-based multilevel analysis. BMC Cancer 2012; 12: 351.

54. Hershman DL, Richards CA, Kalinsky K, Wilde ET, Lu YS, Ascherman JA et al. Influence of health insurance, hospital factors and physician volume on receipt of immediate post-mastectomy reconstruction in women with invasive and non-invasive breast cancer. Breast Cancer Res Treat 2012; 136(2): 535-545.

55. Gort M, Broekhuis M, Otter R, Klazinga NS. Improvement of best practice in early breast cancer: actionable surgeon and hospital factors. Breast Cancer Res Treat 2007; 102(2): 219-226.

56. McDermott AM, Wall DM, Waters PS, Cheung S, Sibbering M, Horgan K et al. Surgeon and breast unit volume-outcome relationships in breast cancer surgery and treatment. Ann Surg 2013; 258(5): 808-814.

57. Kong AL, Pezzin LE, Nattinger AB. Identifying patterns of breast cancer care provided at high-volume hospitals: a classification and regression tree analysis. Breast Cancer Res Treat 2015; 153(3): 689-698.

58. Fisher CS, Martin-Dunlap T, Ruppel MB, Gao F, Atkins J, Margenthaler JA. Fear of recurrence and perceived survival benefit are primary motivators for choosing mastectomy over breast-conservation therapy regardless of age. Ann Surg Oncol 2012; 19(10): 3246-3250.

59. Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. Mammachirurgie: Bundesauswertung zum Erfassungsjahr 2018; Qualitätsindikatoren und Kennzahlen [online]. 23.07.2019 [Accessed: 26.11.2019]. URL: https://iqtig.org/downloads/auswertung/2018/18n1mamma/QSKH_18n1-MAMMA_2018_BUAW_V02_2019-07-23.pdf.

60. Yen TWF, Laud PW, Pezzin LE, McGinley EL, Wozniak E, Sparapani R et al. Prevalence and consequences of axillary lymph node dissection in the era of sentinel lymph node biopsy for breast cancer. Med Care 2018; 56(1): 78-84.

61. Wilson AR, Marotti L, Bianchi S, Biganzoli L, Claassen S, Decker T et al. The requirements of a specialist Breast Centre. Eur J Cancer 2013; 49(17): 3579-3587.

62. Biganzoli L, Marotti L, Hart CD, Cataliotti L, Cutuli B, Kuhn T et al. Quality indicators in breast cancer care: an update from the EUSOMA working group. Eur J Cancer 2017; 86: 59-81.

63. Schrodi S, Tillack A, Niedostatek A, Werner C, Schubert-Fritschle G, Engel J. No survival benefit for patients with treatment in certified breast centers: a population-based evaluation of German cancer registry data. Breast Journal 2015; 21(5): 490-500.

64. Geraedts M, Malik M, Jung O, De Cruppe W. Breast cancer centres in North Rhine-Westphalia: case volume trends 2004–2010. Gesundheitswesen 2013; 75(7): 424-429.

65. Institut National du Cancer. Les autorisations de traitement du cancer [online]. 07.07.2018 [Accessed: 21.10.2019]. URL: https://www.e-cancer.fr/Professionnels-de-sante/L-organisation-de-l-offre-de-soins/Traitements-du-cancer-les-etablissements-autorises/Les-autorisations-de-traitement-du-cancer.

66. Rococo E, Mazouni C, Or Z, Mobillion V, Koon Sun Pat M, Bonastre J. Variation in rates of breast cancer surgery: a national analysis based on French hospital episode statistics. Eur J Surg Oncol 2016; 42(1): 51-58.

The full report (German version) is published under

https://www.iqwig.de/en/projects-results/projects/health-care/v18-05-relationship-between-volumes-of-services-and-quality-of-treatment-outcome-in-the-surgical-treatment-of-breast-cancer-rapid-report.11730.html.

Appendix A – Search strategies

1. MEDLINE

Search interface: Ovid

- Ovid MEDLINE(R) 1946 to May Week 2 2019,
- Ovid MEDLINE(R) Daily Update May 17, 2019

#	Searches
1	Breast Neoplasms/
2	(breast adj3 cancer*).ab,ti.
3	or/1-2
4	((minim* or high* or low or patient or outcome* or importance*) adj3 (volume* or caseload)).ab,ti.
5	((hospital* or center* or centre* or unit* or surgeon* or provider* or physician*) adj2 (factor* or effect*)).ab,ti.
6	((hospital* or center* or centre* or unit*) adj5 (type or level or small* or size)).ab,ti.
7	((hospital* or center* or centre* or unit* or surgeon* or surgical* or physician* or provider*) adj2 (volume* or caseload* or experience* or characteristic* or performance*)).ab,ti.
8	((improve* adj2 outcome*) and (hospital* or center* or centre* or unit* or surgeon*)).ti,ab.
9	((surgeon* or surgical* or physician* or provider* or specialist*) adj3 outcome*).ti,ab.
10	(referral* adj3 (selective* or volume* or rate*)).ti,ab.
11	or/4-10
12	and/3,11
13	12 not (exp animals/ not humans.sh.)
14	13 not (comment or editorial).pt.
15	..l/ 14 yr=2000-Current

Search interface: Ovid

- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations 1946 to May 17, 2019,
- Ovid MEDLINE(R) Epub Ahead of Print May 17, 2019

#	Searches
1	(breast and cancer*).ab,ti.
2	((minim* or high* or low or patient or outcome* or importance*) adj3 (volume* or caseload)).ab,ti.

#	**Searches**
3	((hospital* or center* or centre* or unit* or surgeon* or provider* or physician*) adj2 (factor* or effect*)).ab,ti.
4	((hospital* or center* or centre* or unit*) adj5 (type or level or small* or size)).ab,ti.
5	((hospital* or center* or centre* or unit* or surgeon* or surgical* or physician* or provider*) adj2 (volume* or caseload* or experience* or characteristic* or performance*)).ab,ti.
6	((improve* adj2 outcome*) and (hospital* or center* or centre* or unit* or surgeon*)).ti,ab.
7	((surgeon* or surgical* or physician* or provider* or specialist*) adj3 outcome*).ti,ab.
8	(referral* adj3 (selective* or volume* or rate*)).ti,ab.
9	or/2-8
10	and/1,9
11	10 not (exp animals/ not humans.sh.)
12	11 not (comment or editorial).pt.
13	..l/ 12 yr=2000-Current

2. Embase

Search interface: Ovid

- Embase 1974 to 2019 May 17

#	**Searches**
1	exp Breast cancer/
2	(breast adj3 cancer*).ab,ti.
3	or/1-2
4	((minim* or high* or low or patient or outcome* or importance*) adj3 (volume* or caseload)).ab,ti.
5	((hospital* or center* or centre* or unit* or surgeon* or provider* or physician*) adj2 (factor* or effect*)).ab,ti.
6	((hospital* or center* or centre* or unit*) adj5 (type or level or small* or size)).ab,ti.
7	((hospital* or center* or centre* or unit* or surgeon* or surgical* or physician* or provider*) adj2 (volume* or caseload* or experience* or characteristic* or performance*)).ab,ti.
8	((improve* adj2 outcome*) and (hospital* or center* or centre* or unit* or surgeon*)).ti,ab.
9	((surgeon* or surgical* or physician* or provider* or specialist*) adj3 outcome*).ti,ab.
10	(referral* adj3 (selective* or volume* or rate*)).ti,ab.

#	Searches
11	or/4-10
12	and/3,11
13	12 not medline.cr.
14	13 not (exp animal/ not exp human/)
15	14 not (Conference Abstract or Conference Review or Editorial).pt.
16	..l/ 15 yr=2000-Current

3. The Cochrane Library

Search interface: Wiley

- Cochrane Database of Systematic Reviews: Issue 5 of 12, May 2019
- Cochrane Central Register of Controlled Trials: Issue 5 of 12, May 2019

ID	Search
#1	[mh ^"Breast Neoplasms"]
#2	(breast NEAR/3 cancer*):ti,ab
#3	#1 or #2
#4	((minim* or high* or low or patient or outcome* or importance*) NEAR/3 (volume* or caseload)):ti,ab
#5	((hospital* or center* or centre* or unit* or surgeon* or provider* or physician*) NEAR/2 (factor* or effect*)):ti,ab
#6	((hospital* or center* or centre* or unit*) NEAR/5 (type or level or small* or size)):ti,ab
#7	((hospital* or center* or centre* or unit* or surgeon* or surgical* or physician* or provider*) NEAR/2 (volume* or caseload* or experience* or characteristic* or performance*)):ti,ab
#8	((improve* NEAR/2 outcome*) and (hospital* or center* or centre* or unit* or surgeon*)):ti,ab
#9	((surgeon* or surgical* or physician* or provider* or specialist*) NEAR/3 outcome*):ti,ab
#10	(referral* NEAR/3 (selective* or volume* or rate*)):ti,ab
#11	#4 or #5 or #6 or #7 or #8 or #9 or #10
#12	#3 and #11 with Cochrane Library publication date Between Jan 2000 and Dec 2019, in Cochrane Reviews, Cochrane Protocols
#13	#3 and #11 with Cochrane Library publication date Between Jan 2000 and Dec 2019, in Trials