The volume of surgical interventions and its impact on the outcome: feasibility study based on Belgian data

KCE reports 113C

Federaal Kenniscentrum voor de Gezondheidszorg
Centre fédéral d'expertise des soins de santé
Belgian Health Care Knowledge Centre
2009

The Belgian Health Care Knowledge Centre

Introduction : The Belgian Health Care Knowledge Centre (KCE) is an organization of public interest, created on the 24th of December 2002 under the supervision of the Minister of Public Health and Social Affairs.
KCE is in charge of conducting studies that support the political decision making on health care and health insurance.

Administrative Council

Actual Members : Gillet Pierre (President), Cuypers Dirk (Deputy President), Avontroodt Yolande, De Cock Jo (Deputy President), Baeyens Jean-Pierre, De Ridder Henri, de Stexhe Olivier, Godin Jean-Noël, Goyens Floris, Maes Jef, Mertens Pascal, Mertens Raf, Moens Marc, Perl François, Van Massenhove Frank (Deputy President), Degadt Peter, Verertbruggen Patrick, Schetgen Marco, Devos Daniël, Smeets Yves.

Substitute Members : Cuypers Rita, Decoster Christiaan, Collin Benoit, Stamatakis Lambert, Vermeyen Karel, Kesteloot Katrien, Ooghe Bart, Lernoux Frederic, Vanderstappen Anne, Palsterman Paul, Messiaen Geert, Remacle Anne, Lemye Roland, Poncé Annick, Smiets Pierre, Bertels Jan, Lucet Catherine.

Government commissioner : Roger Yves

Management

Chief Executive Officer a.i. : Jean-Pierre Closon

Information

Federaal Kenniscentrum voor de gezondheidszorg - Centre fédéral d'expertise des soins de santé –
Belgian Health Care Knowlegde Centre.
Centre Administratif Botanique, Doorbuilding (10th floor)
Boulevard du Jardin Botanique 55
B-1000 Brussels
Belgium
Tel: +32 [0]2 287 33 88
Fax: +32 [0]2 287 33 85
Email : info@kce.fgov.be
Web : http://www.kce.fgov.be

The volume of surgical interventions and its impact on the outcome: feasibility study based on Belgian data

KCE reports 113C

FRANCE VRIJENS, KRISTEL DE GAUQUIER, CÉCILE CAMBERLIN

Federaal Kenniscentrum voor de Gezondheidszorg
Centre fédéral d'expertise des soins de santé
Belgian Health Care Knowledge Centre
2009

KCE reports 113C

Title :	The volume of surgical interventions and its impact on the outcome: feasibility study based on Belgian data
Authors :	France Vrijens, Kristel De Gauquier, Cécile Camberlin
Reviewers:	Hans Van Brabandt, Joan Vlayen
External experts :	Xavier de Béthune (Alliance Nationale des Mutualités Chrétiennes), Daniel De Coninck (AZ Sint-Lucas Brugge), Patrick Haentjens (Universitair Ziekenhuis Brussel), Geert Molenberghs (Interuniversity Institute for Biostatistics and statistical Bioinformatics), Victor Legrand (Centre Hospitalier Universitaire de Liège), Antoon Lerut (Universitair Ziekenhuis Leuven), Geert Page (Jan Yperman Ziekenhuis), Pierre Scalliet (Cliniques universitaires Saint-Luc), Luc Renson (Regionaal Ziekenhuis Sint-Trudo), Ward Rommel (Vlaamse Liga tegen Kanker), Yves Taeymans (Universitair Ziekenhuis Gent), Simon Van Belle (Universitair Ziekenhuis Gent), Paul Van Cangh (Cliniques universitaires Saint-Luc), Elizabeth Van Eycken (Belgian Cancer Registry)
Acknowledgements	Stephan Devriese, Kris Henau (Belgian Cancer Registry), Stefaan Van de Sande, Martine Verstreken (Belgian Cancer Registry)
External validators :	Catherine Legrand (Institut de statistique, Université catholique de Louvain), Dirk Schrijvers (Ziekenhuisnetwerk Antwerpen - Middelheim), Paul Sergeant (Universitair Ziekenhuis Leuven)
Conflict of interest :	Dr. Lerut received fees for attending a symposium. Dr. Taeymans received grants from private firms for conducting research. Dr.Renson was remunerated to participate to symposia on total knee replacement. Dr. Scalliet is member of the Board of Directors of IBA which is a private company situated at Louvain-la-Neuve that is active in the fields of cancer diagnosis and therapy.
Disclaimer:	The external experts were consulted about a (preliminary) version of the scientific report. Subsequently, a (final) version was submitted to the validators. The validation of the report results from a consensus or a voting process between the validators. Only KCE is responsible for errors or omissions that could persist. The policy recommendations are also under the full responsibility of KCE.
Layout :	Ine Verhulst

Brussels, 13th July 2009

Study nr 2007-02

Domain: Good Clinical Practice (GCP)

MeSH: Outcome Assessment (Health Care) ; Health Resources ; Neoplasms /surgery ; Cardiovascular Surgical Procedures ; Orthopedics/surgery

NLM classification: W 84.4

Language: English

Format : Adobe® PDF™ (A4)

Legal depot : D/2009/10.273/35

How to refer to this document?

Vrijens F, De Gauquier K, Camberlin C. The volume of surgical interventions and its impact on the outcome: feasibility study based on Belgian data. Good Clinical Practice (GCP). Brussels: Belgian Health Care Knowledge Centre (KCE); 2009. KCE reports 113C (D/2009/10.273/35).

Executive summary

INTRODUCTION AND RESEARCH QUESTIONS

Measuring and understanding the association between the volume and the outcomes of a range of surgical procedures has been the focus of much research since the 1980s. Many of these studies have shown that, for specific procedures, patients admitted to low-volume hospitals or treated by low-volume surgeons have worse outcomes (higher mortality rate, higher complication rate, higher rate of readmissions, etc…) than patients admitted to high-volume hospitals or treated by high-volume surgeons. Until now, very few Belgian studies have been performed on this subject, probably due to the difficult access to the required data.

This study has been designed as a feasibility study, and aims to answer three research questions:

1. For which surgical procedures was the relationship between volume and outcome assessed in the literature? What are the results, and which of these procedures can be studied based on Belgian administrative data?
2. What are the different statistical methods used to assess the volume-outcome relationships?
3. Is it possible to apply these methods on a selection of procedures with Belgian data? What are the pitfalls?

LITERATURE SEARCH AND SELECTION OF PROCEDURES

The first step consisted of a search for good-quality systematic reviews (SR) on the volume outcome association (VOA). Quality assessment was based on the checklist of the Dutch Cochrane Centre which was complemented with an extra evaluation of adjustment for case mix (to take into account the fact that small volume centres or surgeons could treat different patients than high volume ones). The literature search resulted in 25 systematic reviews of either fair or good quality. Altogether, these reviews investigated the VOA for 65 procedures and medical conditions: 13 cardiovascular, 5 neurological, 5 orthopaedic, 1 trauma, 3 intensive care, 18 oncological, 7 gastrointestinal, 6 medical, and 7 miscellaneous.

In a second step, we assessed which of these procedures could be analyzed with Belgian administrative data from the year 2004. Half of the procedures were labelled unfeasible in terms of analysis because they are not available in the Minimal Clinical and Financial Dataset (MCD/MFD) or because their outcomes could not be identified in MCD/MFD. We thus based our study on outcomes that can be identified in the administrative databases (such as mortality and hospital readmissions). This second step resulted in a list of 32 surgical procedures or medical conditions. However, to keep the study within reasonable time limits, the VOA analyses were further limited to the following 12 procedures:

- 5 cancer surgery procedures i.e. oesophageal, pancreatic, colon, breast and lung cancer surgery.
- 4 cardiovascular procedures: carotid endarterectomy (CEA) and carotid stenting (CAS), coronary artery bypass graft (CABG) with/without heart valve replacement or repair, and percutaneous coronary intervention (PCI).
- 3 orthopaedic procedures: total hip replacement, total knee replacement and hip fracture surgery.

The third step aimed at summarizing, for these 12 procedures, the results of the systematic reviews. For this purpose, the SRs were subjected to a second quality appraisal (QA) that assessed the applied method of data synthesis and classified the SRs

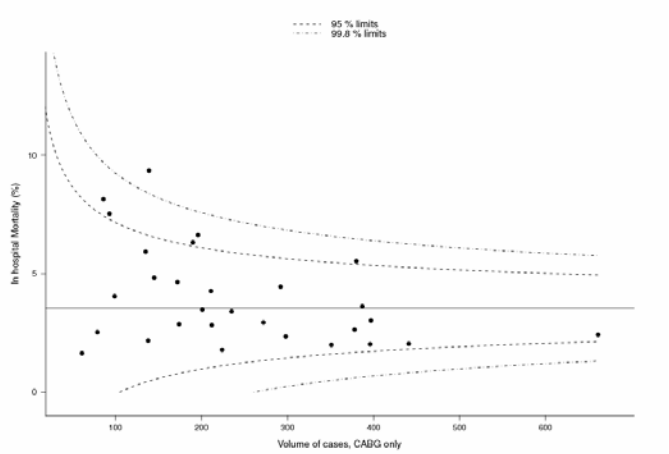
in three groups i.e. very, less and least suitable to support the evidence of a volume outcome association. Finally, the quality of evidence of the primary studies was assessed for each surgical procedure individually. The objective of this third QA was to assign an evidence level to the primary studies used in the systematic reviews according to the GRADE system ranging from high (grade A) over moderate (grade B) to low (grade C).

In general, it has to be emphasised that studies were very heterogeneous in terms of methodological rigor.

CRITICAL REVIEW OF METHODS TO ASSESS VOLUME OUTCOME RELATIONSHIPS

The first methodological question was about the optimal way to present the data graphically. We decided on funnel plots, a standard tool in quality control and meta-analysis that have many advantages when used to compare providers (centres, surgeons, interventionists). Control limits are constructed around the target outcome (the overall mortality, for instance), and define which providers are within or outside expected variability range (taking into account the size of each unit). This plot has the advantage that it draws the reader's eye to important points that lie outside the funnels and that it avoids spurious ranking of institutions (in contrast with league tables). In addition, these plots provide a good basis for an informal graphical check of the volume outcome relationships.

An example of funnel plot: in-hospital mortality after CABG



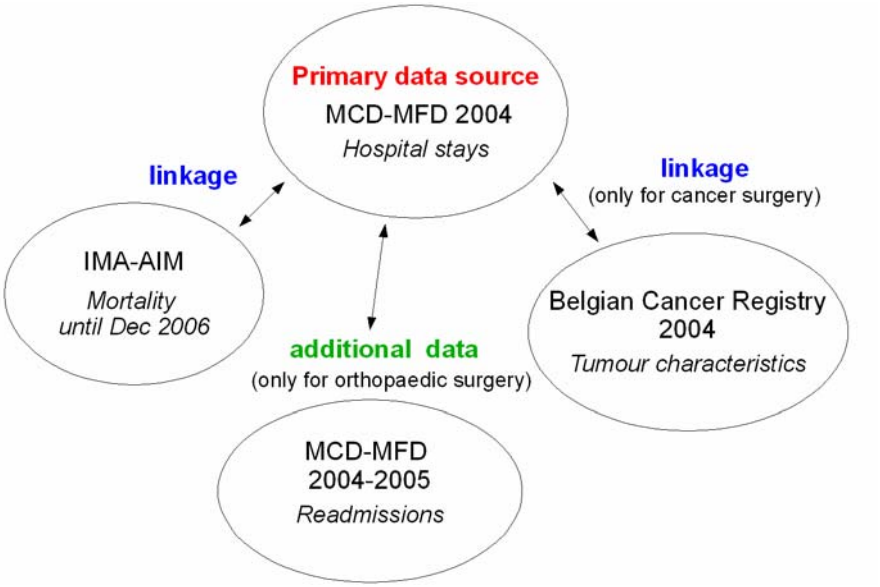
Secondly, one of the major flaws found in the literature was the use of conventional regression models that are based on the assumption that observations are independent of each other. These models ignore the natural hierarchy of data, namely that clusters of patients are operated upon by the same surgeon, and that clusters of surgeons operate in the same hospital. The consequence is that precision of effects is overestimated. A variety of statistical methods exist to produce valid estimations.

Thirdly, with regard to causality, two hypotheses with opposite causal implications have been offered as explanations of the underlying reason for the volume outcome relationships. The first hypothesis is the « learning by doing » or « practice makes perfect » explanation: the more patients are treated by surgeons or hospitals, the better the quality of care. This theory applies to the individual level (experience of surgeon) as well as the hospital level (knowledge transfer, organisation of care). A competing hypothesis is termed « selective referral ». This hypothesis postulates that high quality hospitals, on average, a higher volume, simply because patients, perhaps with the advice of a physician, prefer to be admitted to high quality hospitals Research possibilities on the direction of causality are limited and the statistical models which are used are still being developed.

METHODS APPLIED FOR DATA ANALYSIS OF BELGIAN DATA

SOURCE AND LINKAGE OF DATABASES

Three databases were linked at the patient level. All analyses were performed on recoded data (none of the centres or physicians was identified by name).



DATA DEFINITIONS

Since the initial selection of hospital stays in MCD-MFD 2004 was very elaborate, it was necessary to define, for each procedure, clinically homogeneous patient groups by combining principal diagnosis codes, procedure codes and APR-DRGs.

Volume was defined as the annual number of procedures per hospital and per surgeon. For surgeons operating in different centres, all performed procedures were taken into account.

The following outcomes were analyzed: in-hospital mortality (from MCD data), approximate 30-day mortality (due to a lack of the exact date of death in the IMA data, this outcome covers a post-operative window of at least 1 and maximum 60 days), 2-year mortality (from IMA data), readmission (from MCD data in 2005), complication rates during hospitalisation (from MCD data) and revisions rates at 18 months for hip and knee prosthesis (from MCD data).

STATISTICAL METHODS

Logistic regression models were used to assess the effect of hospital or surgeon volume on the outcome. All effects took into account the correlations of patients within centres and were adjusted for the following case mix parameters: patient age, sex, principal diagnosis of admission, Charlson score (a comorbidity score based on MCD data) and type of procedure (if applicable). For cancer surgery, tumour characteristics were also taken into account (stage, histology and grade).

The analyses were not intended to determine thresholds but were primarily designed to validate the potential existence of volume-outcome correlations.

RESULTS OF DATA ANALYSIS: IS THERE A VOLUME OUTCOME RELATIONSHIP FOR THE SELECTED PROCEDURES?

CANCER SURGERY PROCEDURES

	Oesophagus[1]	Pancreas	Colon[2]	Breast	Lung
Selection of patients					
N patients with cancer diagnosis in Minimal Clinical Data 2004	1 401	1 841	5 756	10 048	7 360
% patients with surgery	27%	17%	48%	86%	18%
N patients included in analysis	342	309	2 724	8 992	1 192
% with info on stage in BCR	67%	67%	72%	79%	75%
Outcomes					
Approximate 30-day mortality	9.1%	8.3%	5.6%	0.25%	3.8%
Two-year mortality	44.7%	53.2%	28.2%	5.2%	26.5%
Association between volume and outcome (2-year mortality) for hospitals					
N hospitals (median volume/yr)	72 (2.5)	74 (2)	114 (21)	114 (56)	97 (7)
Threshold per year (in literature)	6[3]	11[3 & 4]	-	150[8]	-
N hospitals reaching cut off	10	5	-	18	-
Effects in scientific literature?	yes	yes	small	not clear	yes
Effects on Belgian data?	no	yes[10]	no	yes	no
OR 95%CI (low vs high volume) not adjusted for case-mix	0.87 (0.55, 1.38)	1.31 (0.71, 2.42)	0.70[6] (0.44, 1.11)	1.67 (1.36,2.04)	1.03[9] (0.64, 1.64)
OR 95%CI (low vs high volume) adjusted for case-mix	0.82 (0.53, 1.26)	1.29 (0.85, 1.94)	0.64[6] (0.38, 1.07)	1.43 (1.15, 1.79)	1.10[9] (0.67, 1.79)
Association between volume and outcome (2-year mortality) for surgeons					
N surgeons (median volume/yr)	99 (1)	112 (1)	401 (5)	805 (5)	154 (4)
Threshold per year (in literature)	6[3]	-	-	50[9]	-
N surgeons reaching cut off	8	-	-	34	-
Effects in scientific literature?	yes	yes	small	not clear	no
Effects on Belgian data?	yes[10]	yes	yes[10]	yes, small	no
OR 95%CI (low vs high volume) not adjusted for case-mix	1.38 (0.92, 2.08)	1.80[5] (1.06, 3.07)	1.38 [7] (0.82, 2.23)	1.27 (1.00, 1.60)	0.73[9] (0.48, 1.12)
OR 95%CI (low vs high volume) adjusted for case-mix	1.30 (0.88, 1.91)	1.51[5] (1.06, 2.16)	1.36 [7] (0.84, 2.20)	1.25 (0.95, 1.65)	0.82[9] (0.52, 1.28)

[1] including tumours of cardia [2] including recto sigmoid junction [3] US Agency for Healthcare Research and Quality [4] US Leapfrog Group [5] the threshold of oesophagectomies was used because of lack of specific international cut off [6] comparing centres with less than 11/year to centres with more than 60 /year [7] comparing surgeons with less than 6 to more than 20/year [8] Belgian threshold for breast clinics/surgeons since January 2008 [9] using as threshold first quartile (Q1) of 4 interventions per centre and 2 interventions per surgeon [10] not statistically significant, data from several years are needed to obtain precise estimates.
BCR = Belgian Cancer Registry OR 95%CI = Odds ratio 95% confidence interval

Caveat! With respect to the analysis of the VOA for these cancer surgery procedures, following data limitations have to be acknowledged:

Because data were not retrieved or not available in the databases, the following characteristics could not be used for risk adjustment: use of (neo) adjuvant therapy (chemo- or radiation therapy), acuity of admission (elective versus urgent), intention of surgery (palliative versus curative), type of surgical resection (total or partial resection, e.g. lobectomy versus pneumonectomy), menopausal status or oestrogen receptor and progesterone receptor status for breast cancer.

CARDIOVASCULAR PROCEDURES

	CEA[5]	Isolated CABG	Isolated heart valve[3]	PCI
N stays included in analysis	2 860	7 071	1 949	22 561
In-hospital mortality	1.0%	3.5%	5.6%	1.8%
Association between volume and outcome (in-hospital mortality) for hospitals				
N hospitals (median volume/yr)	109 (21)	29 (211)	29(47)	29 (620)
Threshold per year (in literature)	50[1]	200[1]	-	200[1] - 400[2]
N hospitals reaching cut off	11[1]	16[1]	-	28[1] - 21[2]
Effects in scientific literature?	yes	yes	no	only for primary PCI[4]
Effects on Belgian data?	no	yes	no	no, effect reduced with risk adjustment but disease severity is not available
OR 95%CI (low vs high volume) not adjusted for case-mix	1.28[6] (0.45, 3.61)	1.80 (1.34, 2.41)	-	2.20[2] (1.69, 2.86)
OR 95%CI (low vs high volume) adjusted for case-mix	1.38[6] (0.45, 4.27)	1.77 (1.24, 2.53)	-	1.35[2] (0.99, 1.84)
Association between volume and outcome (in-hospital mortality) for surgeons or interventionists (in case of PCI)				
N surgeons (median surgeon volume/yr)	236 (9)	100 (56.5)	92 (14)	215 (51)
Effects in scientific literature?	yes	yes	no	only for primary PCI[4]
Effects on Belgian data?	not enough events to analyze	yes, but smaller than centre effect	not analyzed	no, effect reduced with risk adjustment but disease severity is not available
OR 95%CI (low vs high volume) not adjusted for case-mix	-	-	-	1.69 (1.22, 2.36)
OR 95%CI (low vs high volume) adjusted for case-mix	-	-	-	1.23 (0.92, 1.65)

[1] US Agency for Healthcare Research and Quality [2] US Leapfrog Group [3] Replacement or repair [4] Primary PCI is a PCI which is performed on patients with an acute myocardial infarction, within only a few hours after onset of the symptoms of the AMI [5] Note that the identification of CAS procedures is difficult because of lack of a specific procedure code. Therefore, the number of CAS is probably not complete in this study [6] Effects are reduced when one high volume centre is excluded from analysis.

Caveat data limitations for these cardiovascular procedures:

For CEA and PCI, where death is a rare outcome, complication rate (i.e. acute myocardial infarction or stroke rate) would be a better end point. The difficulty, however, is that the coding of these specific complications in the MCD is unreliable.

Some procedures are not identifiable in the MCD and can therefore not be taken into account in the case-mix adjustment. This limitation applies to the type of stent implanted during PCI (i.e. drug eluting stent, bare metal stent or balloon angioplasty) and to the off pump CABGs.

With respect to cardiac surgery, the EuroSCORE (the most frequently used risk profile system in cardiology) would be better for risk adjustment than the applied Charlson score (which takes into account only comorbidities and not patient's cardiac status). These clinical parameters are, however, not encoded in the MCD.

ORTHOPAEDIC PROCEDURES

	Elective total hip replacement	Elective total knee replacement	Hip fracture surgery
N patients included in analysis	11 856	11 017	9 934
In-hospital mortality	0.24%	0.15%	6.5%
90-day readmission rate for complication	3.3%	3.2%	-
Revision rate at 18 (hip) or 12 (knee) months	1.8%	1.0%	-
Association between volume and outcome for hospitals			
N hospitals (median volume/yr)	115 (84)	114 (78)	113 (78)
Effects in scientific literature?	lack of good-quality evidence	conflicting evidence	yes
Effects on Belgian data?			
90-day readmission rate for complication	yes, small -1.6%[a] (-3.4%, 0.17%)	no -0.4%[a] (-3.0%, 2.3%)	-
Revision rate	no 1.14[b & d] (0.70, 1.85)	no -1%[a] (-3.7%, 1.7%)	-
In-hospital mortality	-	-	no 0.6%[a] (-1.2%, 2.5%)
Association between volume and outcome for surgeons			
N surgeons (median volume/yr)	522 (11)	488 (12)	675 (12)
Effects in scientific literature?	lack of good-quality evidence	conflicting evidence	yes
Effects on Belgian data?			
90-day readmission rate for complication	yes, small -2.7%[a] (-3.9%, -1.4%)	yes, small -2.2%[a] (-3.8, -0.5%)	-
Revision rate	no 1.21[c & d] (0.64, 2.29)	yes -3.7%[a] (-5.9%, -1.4%)	-
In-hospital mortality	-	-	no -0.7%[a] (-2.1%, 0.7%)

[a] Effect on odds of death of a 10% increase in volume [b] Comparing low volume (<60 THR/year) to high volume (>110 THR /year) centres [c] Comparing low volume surgeons (<6 THR/year) to high volume surgeons (>20 THR/year) [d] Hazard ratio and 95% CI.
All results are adjusted for age, sex and Charlson score.

Caveat data limitations for these orthopaedic procedures:

Unfortunately, the MCD do not provide information on the outcomes of greatest interest to patients such as loss of independence, loss of mobility or residual pain.

POSSUM (Physiological and Operative Severity Score for enUmeration of Mortality and morbidity) would be better for risk adjustment in orthopaedic surgery than the applied Charlson score. However, these clinical parameters are not encoded in the MCD.

DISCUSSION AND CONCLUSIONS

This KCE study is a pioneer study for Belgium. Its objective was to answer the following question: can Belgian administrative hospital data be used to study the association between provider volume and outcome? This explains why it was first of all designed as a feasibility study, without the intention of determining volume thresholds, and why it was based on limited number of surgical procedures.

Anyway, the answer to the above mentioned objective is cautiously positive: this study shows that Belgian administrative hospital data can be used to study the volume outcome relationship provided all available information is retrieved from the databases, and, preferably, linked to clinical registries. Linkage with clinical data would be only one way to improve the risk adjustment.

The main **limitation of administrative data** is the impossibility to define the **outcome of interest** for each procedure. Many procedures that were studied in the literature could not be analyzed with our data because of the lack of information on the outcome. Examples include incontinence and quality of life after transurethral prostatectomy or abdominal adhesions in women who had a caesarean section.

For those procedures where it is possible to define the outcome of interest, several precautions should be taken during the analysis:

1. Great care is needed in identifying the **study population** in the administrative databases. Surgical procedures are coded with two different coding systems (ICD-9 classification in the MCD and the nomenclature billing codes in the MFD). There is no 1:1 equivalence between these two coding systems which implies that both codes are needed. In addition, the reason for intervention (i.e. principal diagnosis in MCD) is also necessary to include or exclude specific groups of patients.
2. Serious thoughts must be given to the **time horizon**, i.e. the time between the intervention and the evaluation of the outcome. For complex and therefore high risk procedures such as oesophageal cancer surgery or CABG, outcomes can be assessed at short term (in-hospital, 3-months or 6–months mortality). For less complex procedures or conditions with a good prognosis, such as breast cancer, outcome cannot be assessed at short term simply because there are not enough events. In these cases, evaluation has to be performed in a longer-term perspective, keeping in mind that other treatments besides surgery affect patient's survival.
3. It is important to **distinguish between the effect of the surgeon volume (experience) and the effect of the hospital volume (organisation of care in the broad sense)**. The relative importance of surgeon or hospital volume is difficult to distinguish for infrequent interventions where surgeon volume equals hospital volume.
4. It is important to dispose of **robust information on case mix**. Patient characteristics and disease severity should be available. Risk adjustment is an important issue in volume outcome research because patients with severe co morbidity may be unequally distributed between providers of low and high volume.
    a. MCD data provide information on patient demographics and co morbidities. A useful tool is the Charlson score, which has been validated to predict 1-year mortality. This score can be computed based on MCD data, but consequently inherits the limitations intrinsic to MCD: the quality of the Charlson score depends on the quality and completeness of the coding of co morbidities in each hospital.

b. The second component of the case mix adjustment, i.e. information on disease severity, is not available in MCD data but can be retrieved from existing registries, which record detailed clinical information. Current Belgian registries include the Belgian Cancer Registry and the registry of the Belgian Working Group for Interventional Cardiology. In addition, for hip and knee prostheses, the National Institute for Health and Disability Insurance (NIHDI) recently started the electronic registry ORTHOpride (ORTHOpedic Prosthesis Identification Data).

We cannot rule out confounding by unmeasured characteristics of patients in our study. Nevertheless, we do not believe that limitations related to risk adjustment threaten our main conclusions about the association between volume and outcome.

5. **Appropriate statistical methods** should be used. Regression models are available that respect the hierarchical nature of the data (patients nested within surgeons, surgeons nested within centres), and that account for the correlations within these clusters. The failure to include any type of adjustment for those correlations would lead to falsely high statistically significant effects.
6. The **funnel plot** is a good and "easy to use" tool to present the results graphically. It avoids spurious ranking of institutions, spurious stigmatization of low volume centres, and allows for an informal assessment of any volume outcome relationship.
7. Results should be transparent. In this study, effects of volume were always presented with and without adjustment for case mix. **Sensitivity analyses** are also recommended to test the robustness of the results .
8. We noticed that many hospitals – low-volume as well as high-volume – **missed data on stage** of the cancer and that the percentage of missing data varied among these hospitals. On average, 30% of data on stage was missing. This problem of missing data on disease stage (and other variables useful for risk adjustment) supports the need for complete and accurate data collection.
9. **Sample size** is sometimes not sufficient in one year: analysis of several years for rare tumours or procedures (pancreas, oesophagus) is required.

The fact that this study is based on relatively old data (2004) calls for some caution in the effects observed. Many changes have been introduced in our health care system since 2004: fusion of hospitals, creation of Belgian cancer registry, introduction of the multidisciplinary consultations in oncology, introduction of minimal volumes in the treatment of breast cancer, etc… All those factors have a potential impact on the quality of care. All our results are thus valid for the year 2004, but might not be applicable now.

Our final conclusion is that Belgian administrative data linked to registry data can be an adequate tool to study the volume outcome relationship, provided necessary precautions are taken.

RECOMMENDATIONS

AVAILABILITY AND QUALITY OF ADMINISTRATIVE DATA

- Lack of detailed dates of hospital admission and decease in administrative data hinders precise short-term outcome analysis and benchmarking with international results. The KCE advises the Technical Cell and the IMA to provide complete dates of these events.
- The delay in availability of administrative and registry data is so long that they are no longer representative for the current situation. In June 2009, for example, coupled MCD-MFD data are available for the year 2006 and cancer registry data for 2005. The KCE advises to make administrative data available much sooner..
- The quality and completeness of coding of procedures and comorbidities in MCD is crucial for this type of exercise, but this information is not available. The KCE advises the Ministry of Health to provide more information on timing and audit results of the MCD coding.
- Some procedures are difficult to identify in the MCD data: carotid stenting, drug eluting stents for non-diabetic patients, off pump CABG. The KCE recommends the Ministry of Health to provide clear guidance to hospitals with respect to which ICD-procedure codes should be used in these cases.

IMPROVING OUTCOME OF CARE THROUGH FEEDBACKS AND AUDITING

- The KCE recommends providing feedback on the outcomes of surgical procedures to hospitals and physicians by means of periodic reports. This way, care providers are given the opportunity to compare their own results with those of hospitals and colleagues with a similar patient population. At national level, this reporting can be done to the Colleges of physicians. As a result, the College of Oncology, for example, would be helped in performing its task concerning the evaluation of oncological activities.
- The organisation of such feedbacks would require an institution with expertise in linkage, content and analysis of data. Based on their know how in this domain and their authorisation from the Privacy Commission to link their databases, the KCE proposes giving the responsibility for the feedbacks of oncological procedures to the Belgian Cancer Registry. For non cancer procedures, the KCE recommends to make all existing registries uniform in order to facilitate their management and exploitation.
- The KCE advises the use of funnel plots for the identification of centres that would require further auditing. Audits in potentially problematic centres should be part of a quality control program. This is the responsibility of the Belgian communities or regions.

CANCER SURGERY

A distinction is made on basis of complexity of surgery and incidence of cancer.

- For pancreas cancer surgery, a complex procedure not frequently performed with a high mortality risk, there is abundant evidence that results are better in high volume centres. Therefore, the KCE recommends centralising the expertise in a limited number of centres by establishing an annual minimum threshold of pancreatectomies as it was also recommended in KCE report 105.
- For oesophageal cancer surgery, another example of complex procedure not frequently performed with a high mortality risk, our data did not provide sufficient evidence to establish a recommendation to centralise expertise. Nevertheless, the clear consensus in the scientific literature (which was also reported in KCE report 75) cannot be ignored. Therefore the KCE advises to re-investigate the volume outcome relationship for oesophagectomies for cancer on recent data, including at least two years.
- For breast cancer, our data show that the decision to centralize expertise in breast clinics (starting January 2008) was justified. The KCE recommends to further evaluate the impact of this centralization, including an assessment of other treatments, i.e. radio-, chemo- and hormonal therapy.

CARDIAC SURGERY

- In Belgium, cardiac surgery is performed in 29 so-called B2/B3 centres, with differences in outcomes between small and high volume centres for the coronary bypass surgery (CABG). However, before any recommendation on the use of a minimal volume threshold in cardiac surgery can be made, potential negative effects (e.g. loss of quality due to overburdening) of such a decision have to be studied. At this moment, the KCE advises to identify which processes of care are present in the centres with best outcomes. This way, strategies can be developed to improve care in all centres.
- With respect to percutaneous coronary interventions, the KCE was faced with the limitation that primary PCIs were not identifiable in the MCD 2004. Hence, the VOA for primary PCI, for which evidence was found in literature, could not be assessed in Belgian data. In future studies, however, this problem should be solved since the encoding of STEMI and NSTEMI (Non ST-Elevation Myocardial Infarction) was introduced in the MCD in 2008. The KCE recommends the Ministry of Health to check if hospitals are indeed providing this important information.

ORTHOPAEDIC SURGERY

- Our study showed the difficulties to identify clinical characteristics and outcomes after prosthetic surgery in MCD data. The fact that a registry for knee and hip prostheses has been recently created is good news. The KCE advises to continuously encourage the use of the registry by making registration as easy as possible and facilitating the access to the registered data.

Scientific summary

Table of contents

LIST OF ABBREVIATIONS

Abbreviation	English	Other languages
AAA	Abdominal Aortic Aneurysm	
AHRQ	Agency for Healthcare Research and Quality (USA)	
AIDS	Acquired Immunodeficiency Syndrome	
AMI	Acute Myocardial Infarction	
APR-DRG	All Patient Refined Diagnosis Related Groups	
ASA	American Society of Anesthesiologists	
BCR	Belgian Cancer Registry	Stichting Kankerregister Fondation Registre du Cancer
BMS	Bare Metal Stent	
BWGIC	Belgian Working Group for Interventional Cardiology	
CABG	Coronary Artery Bypass Graft	Coronaire bypass chirurgie Pontage aortocoronarien par greffe
CAS	Carotid Stenting	
CEA	Carotid Endarterectomy	
CI	Confidence Interval	
CVA	Cerebrovascular accident	
DES	Drug-Eluting Stent	
DVT	Deep Venous Thrombosis	
HFS	Hip Fracture Surgery	
IARC	International Agency for Research on Cancer	
ICD-9-CM	International Classification of Diseases, ninth version, clinical modification (WHO)	
IMA-AIM	Common Sickness Funds Agency	Intermutualistisch Agentschap Agence Intermutualiste
IQWiG	German Institute for Quality and Efficiency in Health Care	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KCE	Belgian Health Care Knowledge Centre	Federaal Kenniscentrum voor de Gezondheidzorg Centre fédéral d'expertise des soins de santé
LOS	Length of stay	Verblijfsduur Durée de séjour
MCD	Minimal Clinical Data	Minimale Klinische Gegevens (MKG) Résumé Clinique Minimum (RCM)
MDC	Major Diagnostic Category	
MESH	Medical Subject Headings	
MFD	Minimal Financial Data	Minimale Financiële Gegevens (MFG) Résumé Financier Minimum (RFM)
MoH	Ministry of Health	Federale Overheidsdienst Volksgezondheid, Veiligheid van de Voedselketen en Leefmilieu Service Public Fédéral Santé publique, Sécurité de la Chaîne alimentaire et Environnement
N, Nbr	Number	
NIHDI	National Institute for Health and Disability Insurance	Rijksinstituut voor Ziekte- en Invaliditeitsverzekering (RIZIV) Institut National d'Assurance Maladie-Invalidité (INAMI)

Abbreviation	English	Other languages
NHS	National Health Service	
NS	Not Statistically Significant	
NSTEMI	Non ST-Elevation Myocardial Infarction	
OHT	orthotopic heart transplantation	
OPCAB	Off-Pump CABG	
PCI	Percutaneous Coronary Intervention	Percutane coronaire interventie Intervention coronaire percutanée
Pct	Percentage	
Pctl	Percentile	
PD	Pancreaticoduodenectomy	
PICO	Population Intervention Comparison Outcome	
PTCA	Percutaneous Transluminal Coronary Angioplasty	
SD	Standard deviation	
SMR	Standardised Mortality Rate	
SOI	Severity of illness	
SR(s)	Systematic Review(s)	
SSI	Surgical site infection	
SS	Statistically Significant	
STEMI	ST-Elevation Myocardial Infarction	
TCT	Technical Cell for the processing of hospital data	Technische Cel voor de verwerking van de gegevens met betrekking tot de ziekenhuizen Cellule Technique de traitement de données relatives aux hôpitaux
USA	United States of America	
VOA	Volume Outcome Association	
WHO	World Health Organisation	

1 INTRODUCTION

1.1 GENERAL BACKGROUND OF THE STUDY

Measuring and understanding the association between volume and outcome in surgical procedures has been the focus of much research since the 1980s.[1] Many of these studies have shown that, for specific diagnoses and procedures, patients admitted to low-volume hospitals or treated by low-volume surgeons have a higher mortality rate than patients admitted to high-volume hospitals or treated by high-volume surgeons. There are however many methodological shortcomings: the quality of risk-adjustment techniques varies greatly; very few studies examine the simultaneous contribution of hospital and physician volume to outcomes; the definitions of high and low volume vary so much that the definition of high volume in one study can be the number used to indicate low volume in another study; very few studies explore changes in volume and performance over time; the mathematical nature of this relationship is often only explored as linear , hereby neglecting a possible asymptotic relationship; finally, missing from most studies is an exploration of the mechanism through which volume influences outcome.[1]

1.2 AIMS, SCOPE AND METHODS

This study aims at answering three research questions:

1. For which surgical procedures was the relationship between volume and outcome assessed in the literature? What are the results, and which of these procedures can be studied based on Belgian administrative data?
2. What are the different statistical methods used to assess the volume-outcome relationships?
3. Is it possible to apply these methods on a selection of procedures with Belgian data? What are the pitfalls?

The first step of this study was to identify from the scientific literature a list of diagnoses and interventions for which the relationship between hospital and physician volume and patient outcome has been studied, and that can be analyzed in Belgian data. This search was based on systematic reviews of the literature available to September 2008 (but more recent individual studies were also discussed).

Secondly, an overview was given of the different statistical methods to assess the volume-outcome relationship, with causality as a specific point of interest.

Thirdly, these statistical methods were implemented on a selection of procedures and medical conditions. Data were obtained from the Belgian Minimal Clinical and Minimal Financial Data (MCD and MFD) for the year 2004. These were linked to data from the Common Sickness Funds Agency (IMA-AIM) in order to obtain out-of-hospital mortality and socio-demographic characteristics. A second linkage has been established with clinical data from the Belgian Cancer Registry (BCR), but only for the oncologic procedures.

1.3 REPORT STRUCTURE

The report is composed of 9 chapters including an introduction (Chapter 1).

The main chapters address the following topics:

- Literature search and selection of procedures and conditions (Chapter 2)
- Critical review of methods to assess the volume outcome relationship (Chapter 3)
- Methods applied for the analyses (Chapter 4)
- Results for five cancer surgery procedures (Chapter 5)
- Results for four cardiovascular procedures (Chapter 6)
- Results for three orthopaedic procedures (Chapter 7)

Discussion of these results is provided in Chapter 8. Finally, Chapter 9 provides a list of references used in the report.

2 LITERATURE SEARCH AND SELECTION OF PROCEDURES AND CONDITIONS

2.1 RESEARCH QUESTIONS

The goal of this first chapter is to investigate the relationship between hospital and physician volume and, on the other hand, patient outcome for certain diagnoses and interventions in health care based on a systematic search of the literature available to September 2008.

This objective was reached in three steps, each with a corresponding research question.

1. Which systematic reviews study the relationship between hospital and physician volume and patient outcome? For which surgical procedures and medical conditions did these systematic reviews investigate this relationship?
2. Which of these surgical procedures and medical conditions can be studied with Belgian data? Which outcome variables are selected?
3. What were the results of these systematic reviews in relation to the volume-outcome association (VOA) for the 'Belgian' selection of procedures and conditions?

2.2 METHODOLOGY

2.2.1 Methodology for the selection of systematic reviews (first research question)

The literature search was done according to the KCE Process notes Search for Evidence & Critical Appraisal: Good Clinical Practice (GCP).[2] The search was performed by one reviewer, validated by another KCE expert and, finally, discussed by a group of experts.

The first step consists of a search for evidence synthesis i.e. systematic reviews, meta-analyses, Health Technology Assessment reports and other good-quality, directly relevant evidence synthesis.

2.2.1.1 Search protocol in relation to evidence syntheses

Concerning the search for evidence syntheses, the key components of the search question were patients undergoing a certain procedure or presenting a certain disease (*Population*) in a high volume hospital or by a high volume surgeon (*Intervention*), compared with a low hospital or physician volume (*Comparison*), assessing outcome measures like mortality, length of stay, complication rate or resource utilization costs (*Outcome*). In order to broaden our search we did not limit the population to specific conditions of procedures.

A search was performed of the Medline and Embase databases (1966 until September 2008) and the Cochrane Library. Search algorithms combined following medical subject heading (MeSH) terms and key words: "Volume", "Outcome", "Regionalization", "Quality Indicators", "Length of Stay", "Recovery of Function", "Complications", "Mortality" and "Health Resources". The search was limited to systematic reviews and meta-analyses. We did not use a language restriction. Reference lists of retrieved articles were used to complete our search. We also performed a manual search of websites of HTA Agencies and of the Agency for Healthcare Research and Quality. Details on the search strategy are provided in Appendix 1.

A first selection of articles was based on the title. Papers that were clearly not relevant to the key questions were eliminated. Then a selection was made based on the abstracts. Finally, the same reviewer confirmed the eligibility of the identified studies by reading the entire articles. The inclusion or exclusion decisions based on abstracts and full texts were based on criteria resulting from the PICO components of the search question.

Articles that were not about patients undergoing a certain procedure or presenting a certain condition were excluded. Studies that did not compare high and low volume physicians or hospitals were not included. Narrative reviews, observational studies, economic analyses or other studies that were clearly no reviews were excluded. We also excluded duplicate publications or publications reporting repeatedly on the same study population. In these cases, the most recent or most complete article was retained. Contrary to the absence of a language restriction in the primary search, it was considered appropriate to apply language as an exclusion criterion at this stage. Therefore, articles in languages other than English, German, Dutch or French were excluded.

2.2.1.2 First quality appraisal of systematic reviews

Quality assessment was performed by one reviewer on the basis of the full-text and consisted of two parts. First, we used the checklist of the Dutch Cochrane Centre (www.cochrane.nl) concerning systematic reviews of observational studies. This checklist was translated into English and is provided in Appendix 2.

Since some pre-existing patient factors, such as disease severity and co-morbidities, are clear determinants of outcome, independent of volume, it is necessary to account for differences in such factors to make valid comparisons between high- and low-volume providers.[3] Therefore we considered it important to add the assessment of risk-adjustment to our quality appraisal, complementing the checklist of the Dutch Cochrane Centre with an extra item i.e. number 5.

The overall quality of the articles was summarized as good, fair or poor quality. Systematic reviews received a poor quality appraisal when the search of the literature was insufficient and no risk-adjustment of included studies was reported and/or the review failed to describe the main characteristics of the original studies. Good-quality systematic reviews were flawless i.e. all quality criteria scored well. Finally, we assigned a fair-quality label to those systematic reviews with one or more, admissible shortcomings. Poor quality studies were excluded from further use for this study. Good- and fair-quality systematic reviews, on the other hand, were used to draw up a list of surgical procedures and medical conditions that have been studied in recent years.

2.2.2 Methodology for the selection of procedures, conditions and outcome measures (second research question)

The literature search resulted in a long list of procedures and conditions which were studied in the selected systematic reviews.

Subsequently, it was decided which procedures or conditions of this list would be analysed with Belgian data. Data were provided by three databases. The federal Ministry of Health (MoH) together with the National Institute for Health and Disability Insurance (NIHDI) maintain a database of all Belgian hospital stays, called The Belgian Minimal Clinical and Financial Data (MCD and MFD). These data were linked to data from the Common Sickness Funds Agency (IMA-AIM) in order to obtain out-of-hospital mortality and socio-demographic characteristics. A second linkage was established with clinical data from the Belgian Cancer Registry (BCR), but only for the oncologic procedures. See Figure 4.1 on page 39.

The exclusion of a surgical procedure or medical condition for further analysis in Belgian data is based on following criteria:

1. Hospital stays in relation with this procedure/condition cannot be retrieved from the MCD/MFD.
2. The procedure or treatment is already regionalized in Belgium.
3. In this report, the term regionalization has nothing to do with politics, but refers to a process by which specialized procedures like paediatric heart surgery are deliberately distributed in a presumably rational and efficient geographic context, as Hannan explained it.[4]
4. Procedure or condition is rare.

5. Volume-outcome relation was studied in previous KCE reports.
6. Outcome measure cannot be analyzed with the available data.
7. Procedure/intervention is too heterogeneous or there is an overlap with other procedures.
8. There is not enough information on the stage of the cancers in the BCR data i.e. the number of records with a known 'Combined-stade' is insufficient.

Finally, an assessment was done of which outcome measures were available for analysis.

2.2.3 Methodology to retrieve information on the volume-outcome association from the scientific literature (third research question)

2.2.3.1 Selection of relevant systematic reviews that study selected procedures or conditions

Given that only a limited number of procedures and conditions will be analysed with the help of Belgian data, systematic reviews that exclusively study nonselected procedures were excluded from further use.

2.2.3.2 Second quality appraisal of systematic reviews

The remaining systematic reviews will be subjected to a second quality appraisal that assesses the method of data synthesis that was applied in the systematic review.

Gandjour et al. underlined how problematic the data synthesis of primary studies on the volume outcome relationship can be because of the heterogeneity of studies in terms of patient characteristics, extent of follow up, volume thresholds, and types of diagnoses and interventions.[5] For that reason, this second quality appraisal aims at classifying the systematic reviews in three groups:

- SRs that are labelled as '+++' are very suitable as supporting evidence in relation to the volume outcome association because they use a specific method to identify the studies most likely to provide an unbiased estimate of the effect of volume on outcome.
- SRs with the label '++' are less suitable as supporting evidence. In these SRs, studies are scored or categorized according to the applied risk-adjustment, but, finally, all studies remain eligible for discussion of the VOA.
- SRs with the label '+' are least suitable. Here, there is a certain assessment of the risk-adjustment (although it is not always clear how this assessment is done), but it is not used as a selection criterion.

2.2.3.3 Data extraction methodology

Data extraction is performed for each procedure or condition that was selected to be studied with Belgian data. The evidence tables (see Appendices 8, 9 and 10) are organized by procedure or condition and contain information from all the systematic reviews that remained after the second quality appraisal and that analyse the volume outcome association for this procedure or condition.

The data extraction template is provided in Appendix 3. First, this table contains information on the totality of systematic reviews for the studied surgical procedure or condition: total number of SRs included, publication date of all SRs, total number of studies included, study period of all included studies, and, finally, outcome variables studied in these SRs. Second, each individual SR is described in relation to its main findings, threshold volumes and quality of evidence. On the basis of all this information, conclusions are formulated on the VOA for the procedure concerned.

MAIN FINDINGS

This section summarizes the results of the SR in relation to the volume outcome association. These results are given for each outcome variable, and, when possible, for hospital volume and surgeon volume separately.

For reasons of feasibility we simplified the results into four possible conclusions on this volume outcome association.

1. Direct or positive VOA when volume and outcome move in the same direction i.e. better outcome in high-volume hospitals or high-volume surgeons.
2. Inverse or negative VOA when volume and outcome move in opposite direction i.e. outcome worsens with increasing volume (Note that if outcome is expressed in a negative way i.e. as mortality or complication rate, an inverse VOA means that mortality decreases when volume increases).
3. No association between volume and outcome when there is any change in outcome with increasing volume.
4. Conflicting results on the VOA when some primary studies indicate a positive and others no or a negative relation.

These four categories do not reflect other possibilities of the VOA, for example upper limits of the volume effect (reaching asymptotic in quality or decreasing in quality after a certain upper threshold).

When the VOA is positive or negative, the statistical significance of the primary studies is summarized between brackets. Statistical significance was set at a level of significance of 5%.

- '>75% SS' indicates that three out of four primary studies were statistically significant;
- '>75% NS' means that three out of four primary studies showed a statistically not significant trend toward a better or worse outcome in high-volume hospital or high-volume surgeon;
- 'mix SS & NS' tells the reader that there is a mixture of SS and NS primary studies and that neither of these represent more than 75% of the totality of primary studies.

THRESHOLD VOLUMES

This is a summary of the thresholds for surgeons or hospitals that were defined in the systematic review or, more often, in the primary studies that were included in the SR. The threshold stands for the annual number of procedures or conditions that is performed or treated in the hospital, or by the surgeon.

The threshold can be an exact number. In most cases, however, since several primary studies are summarized, the threshold is presented as a range which means that the defined volume fluctuates between a minimum and a maximum number of procedures.

There are four possible thresholds: for a high volume hospital, a low volume hospital, a high volume surgeon and a low volume surgeon. Because of the divergence in definition of a high-volume and low-volume provider, there will often be an overlap between these two thresholds.

QUALITY OF EVIDENCE

The quality of evidence of the primary studies is assessed for each procedure individually. The objective of this third quality appraisal is to assign an evidence level to the primary studies according to the GRADE system ranging from high (grade A) over moderate (grade B) to low (grade C).[6] In theory, observational studies generally yield only grade C evidence. But Guyatt et al. mention that observational studies with overwhelming evidence can be classified as of grade A quality. In addition, strong evidence from observational studies can be scored as grade B quality.[6]

On the basis of Guyatt's grades of evidence, four criteria were formulated for the assignment of a Grade B or Grade C level of evidence to primary studies that were analysed in systematic reviews. Grade A level of evidence was considered impossible.

1. Fitness of the SRs to support evidence, as assessed in the second quality appraisal (see Appendix 7):
   - Grade B: ++ or +++
   - Grade C: +
2. Number of primary studies in the SR that study the VOA for this procedure or condition:
   - Grade B: ≥ 5 primary studies
   - Grade C: < 5 primary studies
3. Study period of the primary studies:
   - Grade B: in or after 1988
   - Grade C: before 1988
4. Percentage of studies with a positive or negative volume outcome association:
   - Grade B: at least 75% of primary studies show a statistically significant relation (positive or negative) between volume and outcome
   - Grade C: at least 75% of primary studies show a statistically not significant relation (positive or negative) between volume and outcome, or there is a mix of statistically significant and not statistically significant primary studies.

   This last criterion can be criticized as statistical significance depends on size of effects and of sample size. A typical characteristic of volume outcome studies is that sample size depends on the prevalence of the disease or the intervention.

CONCLUSION

The conclusion on the association between volume (i.e. number of procedures or conditions in hospitals or by surgeons) and the different outcomes will preferably be based on systematic reviews with a Grade B level of evidence. Outcome measures, for which there is insufficient evidence, will not be summarized in the conclusions.

In relation to the threshold, an attempt is made to summarize the different thresholds (i.e. high-volume and low-volume) into one single minimal hospital volume threshold, and, if possible, one minimal surgeon volume threshold. Such volume standards seem to make more sense than extensively overlapping thresholds. For some procedures, the thresholds are also based on those issued by organisations such as the US Agency for Healthcare Research and Quality and the US Leapfrog Group. The rationale behind the threshold is explained in the tables of evidence (Appendix 8, 9 and 10).

2.2.3.4 Search protocol in relation to additional primary studies

In chapters 5, 6 and 7 of this report, the results of the analysis of the Belgian data will be compared with the results from the scientific literature. For a limited amount of procedures and conditions, the number of studies that was retrieved through the selected systematic reviews was considered insufficient as comparison material for this discussion. For these procedures and conditions, an additional literature search was performed in an attempt to complement the systematic reviews with more recent primary studies.

Such an additional search was done for the following procedures: heart valve replacement or repair, hip fracture surgery, oesophageal, pancreatic, colon, breast and lung cancer surgery.

The search protocol in relation to these primary studies is very comparable to the one performed for the evidence syntheses which is described on page 7 (paragraph 2.2.1.1), with the singularity that the search focuses on patients who underwent a specific procedure or presented a certain disease. Details on the search strategy are provided in Appendix 4.

2.3 RESULTS

2.3.1 Results in relation to the selection of systematic reviews (first research question)

2.3.1.1 Number of retrieved articles, in- and excluded studies, and reasons for exclusion

From a total of 905 studies identified by our search, 77 were potentially relevant and were all reviewed in detail.

One study was excluded because it was a process outcome evaluation instead of a volume outcome evaluation.[7] One study was excluded because it compared the volume-outcome associations between the USA and Canada.[8] Twenty-nine studies were excluded because they were not systematic reviews. It concerned 19 narrative reviews,[9-27] 1 primary observational study,[28] 1 report on the use of volume as an inpatient quality indicator and,[29] 5 discussion papers about the volume-outcome relationship,[30-34] and finally, two synthetic reviews that gave an overview of systematic overviews.[35, 36] Eight studies were excluded because they reported repeatedly on the same study population.[3, 37-43] One Norwegian and one Spanish study were excluded because of the language.[44, 45]

After this selection, 36 articles remained.

It has to be mentioned that most studies were retrospective and therefore possibly prone to bias.

2.3.1.2 Results of first quality appraisal of systematic reviews

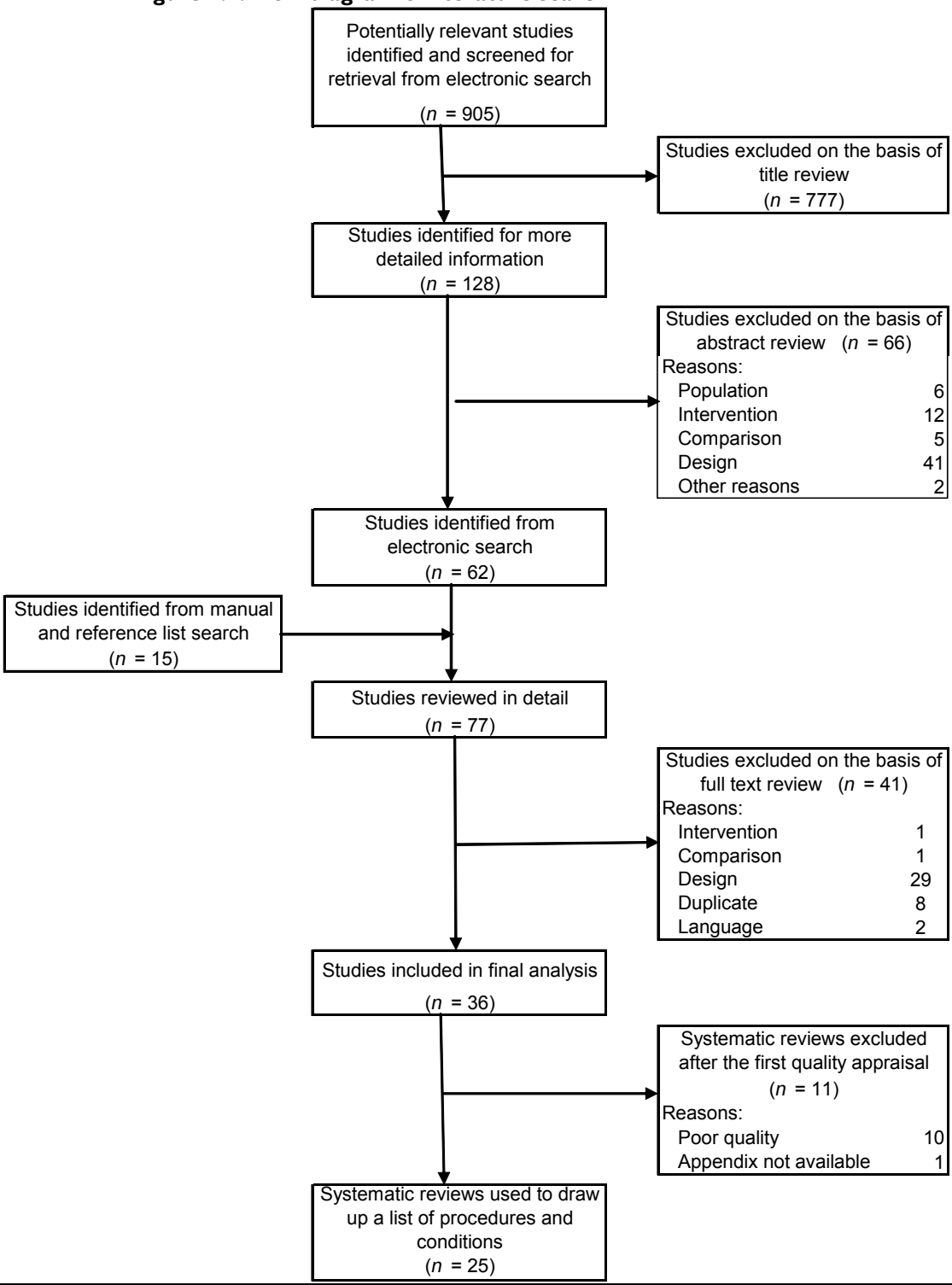
Ten systematic reviews were assessed as poor quality and were excluded.[46-55] The remaining 26 systematic reviews consisted of 13 fair-quality SRs and 13 good-quality SRs. Appendix 5 provides an overview of the critical appraisal.

Unfortunately, the fair-quality SR by Shervin et al. had to be excluded because we could not obtain the appendices of the article, necessary for data extraction.[56] These appendices were not available at the journal's site nor could they be retrieved from the authors.

Finally, 25 systematic reviews were eligible for the drawing up of a list of procedures and conditions.

Figure 2.1 summarizes the literature search.

Figure 2.1: Flow diagram of literature search



2.3.2 Results in relation to the selection of procedures, conditions and outcome measures (second research question)

The literature search resulted in a long list of 65 procedures and conditions which were studied in the selected systematic reviews: 13 cardiovascular, 5 neurological, 5 orthopaedic, 1 trauma, 3 intensive care, 18 oncologic, 7 gastrointestinal, 6 medical, and 7 miscellaneous. All these are listed in Appendix 6.

2.3.2.1 *Selection of procedures and conditions*

Firstly, the exclusion of a surgical procedure or medical condition for further analysis in Belgian data is based on following criteria:

1. Hospital stays in relation with this procedure or condition cannot be retrieved from the Belgian Minimal Clinical and Financial Data (MCD/MFD).
   - First, MCD/MFD of day-care hospital stays cannot be linked to the IMA-AIM data because of lack of correspondence tables between both databases. This implicates that out-of-hospital mortality cannot be obtained for these stays, rendering them unsuitable for analysis. Following procedures and conditions are largely performed in day-care and are therefore excluded from the analysis: cardiac catheterization, cataract surgery and inguinal hernia repair.
   - Second, some procedures or conditions are very difficult to identify in the MCD/MFD. In most volume outcome studies analyzing severe trauma patients, hospital stays are identified through a severity of injury score.[20] Since a similar score is not available in the MCD/MFD, trauma is excluded from analysis.
2. Procedure or treatment is already regionalized (i.e. performed in a limited number of centres) in Belgium.
   - In the areas of paediatric heart surgery, paediatric oncology, placement of ventriculoperitoneal shunt in paediatric surgery and overall care and treatment of human immunodeficiency virus (HIV) or acquired immunodeficiency syndrome (AIDS) the provision of health care services has already been regionalized, which means that these services are already limited to expert centres.
3. Procedure or condition is rare.
   - Surgery for acoustic neurinoma (also called vestibular Schwannoma), pancreatic transplantation are procedures of which there are only a few cases a year in Belgium, which makes them unsuitable for analysis.
4. Volume-outcome relation was studied in previous KCE reports.
   - Acute myocardial infarction has been studied previously.[57]
5. Outcome measure cannot be analyzed with the available data.
   - The out-of-hospital mortality data are available until the year 2006. Readmissions can be investigated until 2005. Several procedures or conditions have outcome measures on the long term, for example, rejection rate after kidney transplantation, amputation rate after lower extremity arterial bypass, incontinence after transurethral prostatectomy, or adhesions in women who had a caesarean section.
6. Procedure or condition is too heterogeneous or there is an overlap with other procedures.
   - Craniotomy, cholecystectomy, intensive care, respiratory insufficiency, COPD, urinary infection, and cirrhosis, all represent very heterogeneous patient populations.
   - Mechanical circulatory support and pelvic cancer surgery are complex interventions which overlap with other procedures.
7. There is not enough information on the stage of the cancers in the BCR data. For the following cancers, for example, the BCR data have only limited information on the stage: C22 Liver and intrahepatic bile ducts (44% of records with known Combined-stade); C56 Ovary (54%); C32 Larynx (53%).[58]

Details on the selection process for each individual procedure or condition are described in Appendix 6.

The initial idea was to study the volume outcome relationship for 32 procedures and medical conditions. To this end, we selected more than 400 000 hospital stays in the MCD 2004. This large amount of data, however, obliged us to limit the IMA data that were linked to MCD; although IMA has information on the use of (neo) adjuvant therapy, only the decease date of patients was retrieved from the IMA database.

Nevertheless, when we were faced with the abundance of 32 procedures and conditions, it was decided to limit the analyses to 12 surgical procedures. These were chosen from 3 medical domains: four cardiovascular procedures, three orthopaedic and five oncologic procedures. The choice of procedures was done in agreement with Belgian Cancer Registry and internal KCE experts.

12 procedures will be analysed in Belgian data:

Four cardiovascular procedures: carotid endarterectomy and carotid stenting; coronary artery bypass graft; heart valve replacement or repair; percutaneous coronary intervention.

Three orthopaedic procedures: total hip replacement; total knee replacement; hip fracture surgery.

Five oncologic procedures: oesophageal cancer surgery; pancreatic cancer surgery; colon cancer surgery; breast cancer surgery and lung cancer surgery.

2.3.2.2 Available outcome measures

The following information is unavailable in the data:

1. Mortality within a very specific time frame i.e. 30-day mortality is impossible since the IMA-AIM mortality date only mentions month and year of decease. As best possible approximation of the 30-day mortality, we included the mortality until the last day of the month following the procedure. For convenience, this is called the approximate 30 day-mortality. The reader should be aware, though, that this outcome measure covers a post-operative window of at most 60 days. Two examples might clarify:
   - If a patient undergoes a CABG on January 1st, 2004 and the data tell us that he deceased by the end of February 2004, we know that he lived at least 30 days but less than 60 days after the CABG.
   - A patient who has a PCI on January 31st, 2004, who is deceased by the end of February 2004, will have lived at least 1 day but less than 30 days after the PCI.
2. Mortality data for this project (provided by IMA-AIM) were only available until the end of 2006. Consequently, mortality of three years or more after a procedure (which took place in 2004) is not available. This implies that two-year survival is the maximum feasible outcome measure for long-term mortality.
3. Complication rates within a specific time frame are unavailable because the MCD admission date only specifies month and year. This makes it, for example, impossible to study the 30-day infection rate.
4. Information about cancer recurrence is not readily available in the BCR data.

Although information on the length of stay and costs of hospital stays is available in the MFD, these measures will not be analysed while this study focuses on clinical endpoints and not on resource utilization.

Table 2.1 gives an overview of the procedures and their outcome variables that were selected to be analysed with Belgian data.

Table 2.1: Procedures and conditions and their outcome variables for which the volume outcome association will be analysed in Belgian data

Procedure or condition (SRs in which studied)	Outcome measures				
	PRIMARY HOSPITALIZATION		READMISSION		OUT-OF HOSPITAL
	In-hospital mortality	Complication rate	Revision rate	Complication rate	Mortality
Cancer surgery					
Oesophageal cancer surgery [1, 5, 59, 60, 62, 64, 69, 70]	x				at 2 years
Pancreatic cancer surgery [1, 5, 59, 60, 62, 64, 69, 71, 72]	x				at 2 years
Colon cancer surgery [1, 5, 59, 60, 62, 64, 69, 73-75]	x				at 2 years
Breast cancer surgery [1, 5, 60, 62, 64, 69]	x				at 2 years
Lung cancer surgery [1, 5, 59, 60, 62, 64, 69]	x				at 2 years
Cardiovascular procedures					
Carotid Endarterectomy [1, 5, 59-65]	x	- Stroke rate - AMI rate			Approximate 30-day mortality
Coronary artery bypass graft [1, 5, 59, 60, 62, 64, 66]	x				
Heart valve replacement and repair [62]	x				
Percutaneous coronary intervention [1, 5, 59, 60, 64, 67]	x	- AMI rate - CABG rate			
Orthopaedic surgery					
	x	- Deep wound infection	x	- Dislocation rate	

Procedure or condition (SRs in which studied)	Outcome measures				
	PRIMARY HOSPITALIZATION		READMISSION		OUT-OF HOSPITAL
	In-hospital mortality	Complication rate	Revision rate	Complication rate	Mortality
Total hip replacement [1, 5, 59, 60, 62, 64]		- Pulmonary embolism - Deep venous thrombosis.	at 18 months	- Deep wound infection - Pulmonary embolism all at 90 days	
Total knee replacement [1, 5, 59, 60, 62, 64, 68]	x	- Deep wound infection - Pulmonary embolism - Deep venous thrombosis.	x at 1 year	- Deep wound infection - Pulmonary embolism - AMI - Pneumonia - Deep venous thrombosis all at 90 days	
Hip fracture surgery [1, 5, 64]	x				

2.3.3 Results in relation to the volume-outcome association (third research question)

2.3.3.1 Selection of relevant systematic reviews that study procedures or conditions which were selected

Systematic reviews that exclusively study nonselected procedures were excluded from further use. This way, we excluded 6 SRs that analyse Abdominal Aortic Aneurysm (AAA) repair or urologic cancers.[76-81]

After this exclusion, nineteen systematic reviews were left.[1, 5, 59-75]

2.3.3.2 Second quality appraisal of systematic reviews

Nine SRs are labelled as '+++' which means that they are very suitable as supporting evidence in relation to the volume outcome association because they use a specific method to identify the studies most likely to provide an unbiased estimate of the effect of volume on outcome.[5, 60, 63, 67, 69, 70, 74, 75] Three SRs get the label '++' and are less suitable as supporting evidence.[61, 64, 65] Finally, there are seven SRs with the label '+' which are least suitable as supporting evidence.[1, 59, 62, 66, 68, 71-73]

Appendix 7 provides an overview of all SRs and their fitness to support evidence in relation to the volume outcome association.

2.3.3.3 Results of the systematic reviews

Detailed results of the literature review, using the data extraction template, are described in the evidence tables in Appendix 8 (cancer surgery procedures), Appendix 9 (cardiovascular procedures) and Appendix 10 (orthopaedic procedures). The conclusions of these evidence tables are repeated in Table 2.2.

During the process of summarizing the systematic reviews into evidence tables, we noticed that many systematic reviews base their results on the same primary studies. We tried to visualize this overlap by listing the primary studies for each systematic review individually, in the evidence tables.

In general, it has to be emphasised that the heterogeneity of studies in terms of methodological rigor is a major problem concerning the systematic reviews as well as primary studies. Despite the quality appraisal of the SRs, there is still a great variability between primary studies in quality of risk-adjustment techniques, patient characteristics, study period, data material (mostly administrative instead of clinical), statistical methods and definitions of high and low volume.

Because of the enormous variability in the definitions for high and low volume, summarizing the thresholds of several primary studies leads to huge overlaps between thresholds of high and low volume in the evidence tables (Appendix 8, 9 and 10). In Table 2.2, however, an attempt is made to summarize the different thresholds into one single minimal hospital volume threshold, and, if possible, one minimal surgeon volume threshold. The rationale behind the summarized threshold is documented in the Appendices 8, 9 and 10.

An in-depth analysis of scientific literature, often at the level of the primary studies, is performed in the discussion sections of this report where scientific literature is compared with the results of Belgian data. The additional primary studies that were retrieved for certain procedures (see Appendix 4) will also be added to this discussion; these studies are unreported anywhere else in the report.

The comparative analyses of scientific literature and Belgian data are to be found in:

- Chapter 5, page 45 for cancer surgery procedures
- Chapter 6, page 128 for cardiovascular procedures
- Chapter 7, page 168 for orthopaedic procedures.

Table 2.2: Summary of finding table on volume - outcome association by procedure

MAIN FINDINGS (on volume outcome association)	THRESHOLD VOLUME	QUALITY OF EVIDENCE
Cancer surgery procedures		
Oesophageal cancer surgery		
INVERSE relation between: - hospital volume – mortality - surgeon volume - mortality	Minimal hospital volume threshold: - threshold 1: 6 oesophagectomies per annum - threshold 2: 13 oesophagectomies per annum.	- 2 out of 8 SRs have Grade B evidence level [69, 70] - 6 out of 8 SRs have Grade C evidence level [1, 5, 59, 60, 62, 64]
Pancreatic cancer surgery		
INVERSE relation between: - hospital volume – mortality - surgeon volume - mortality	Minimal hospital volume threshold: 11 pancreatic resections per annum.	- 3 out of 9 SRs have Grade B evidence level [5, 60, 69] - 6 out of 9 SRs have Grade C evidence level [1, 59, 62, 64, 71, 72]
Colon cancer surgery		
INVERSE relation between: - hospital volume – mortality - surgeon volume - mortality	Threshold not possible to summarize.	- 2 out of 9 SRs have Grade B evidence level [69, 74, 75] - 7 out of 9 SRs have Grade C evidence level [1, 5, 59, 60, 62, 64, 73]
Breast cancer surgery		
Limited evidence does not allow a conclusion on the volume-outcome association for breast cancer surgery.	Minimal hospital volume threshold: - threshold 1: 10 breast cancer surgery procedures per annum - threshold 2: 150 breast cancer surgery procedures per annum.	Six out of six SRs have Grade C evidence level [1, 5, 60, 62, 64, 69] Cave! Very few primary studies for each SR.
Lung cancer surgery		
INVERSE relation between: - hospital volume - mortality	Threshold not possible to summarize.	- 1 out of 7 SRs has Grade B evidence level [69] - 5 out of 7 SRs have Grade C evidence level [1, 5, 59, 60, 62]
Cardiovascular procedures		
Carotid Endarterectomy (CEA) and Carotid Stenting (CAS)		
INVERSE relation between: - hospital volume - mortality - hospital volume - stroke rate - surgeon volume – mortality - surgeon volume – stroke rate	Minimal hospital volume threshold = 79 CEA per annum.	- 5 out of 8 SRs have Grade B evidence level [5, 60, 61, 63, 65] - 3 out of 8 SRs have Grade C evidence level [1, 59, 62]
Coronary artery bypass graft (CABG)		

MAIN FINDINGS (on volume outcome association)	THRESHOLD VOLUME	QUALITY OF EVIDENCE
INVERSE relation between: - hospital volume - mortality - surgeon volume – mortality	Minimal hospital volume threshold = 200 CABG per annum.	- 1 out of 7 SRs has Grade B evidence level [60] - 6 out of 7 SRs have Grade C evidence level [1, 5, 59, 62, 64, 66]
Heart valve replacement or repair		
Limited evidence does not allow a conclusion on the VOA for heart valve replacement or repair.	No threshold available.	Only 1 SR with Grade C evidence level. [62]
Percutaneous coronary intervention (PCI)		
CONFLICTING results (i.e. mix of positive and negative relation) for: - hospital volume - mortality for PCI for mixed indications - operator volume – mortality for PCI for mixed indications - hospital volume – emergency CABG rate INVERSE relation between: - hospital volume – mortality for primary PCI - operator volume – mortality for primary PCI - operator volume - emergency CABG rate	Minimal hospital volume threshold: - threshold 1: 200 PCI per annum - threshold 2: 400 PCI per annum.	- 3 out of 6 SRs have Grade B evidence level [5, 60, 67] - 3 out of 6 SRs have Grade C evidence level [1, 59, 64]
Orthopaedic procedures		
Total hip arthroplasty/replacement (THR)		
CONFLICTING results (i.e. mix of positive, negative and no relation) for: - hospital volume – in-hospital mortality	Minimal hospital volume threshold: - threshold 1: 10 THR per annum - threshold 2: 100 THR per annum.	- 1 out of 5 SRs has Grade B evidence level [60] - 4 out of 5 SRs have Grade C evidence level [1, 5, 59, 62]
Total knee arthroplasty/replacement (TKR)		
INVERSE relation between: - hospital volume - mortality CONFLICTING results (i.e. mix of positive and negative relation) for: hospital volume – post-operative complication rate	Minimal hospital volume threshold: - threshold 1: 10 TKR per annum - threshold:2 100 TKR per annum.	- 1 SR has Grade B evidence level [68] - 6 out of 7 SRs have Grade C evidence level [1, 5, 59, 60, 62, 64]
Hip fracture surgery		
INVERSE relation between: - hospital volume - mortality	Threshold not possible to summarize.	- 3 out of 3 SRs have Grade C evidence level [1, 5, 64]

3 CRITICAL REVIEW OF METHODS TO ASSESS THE VOLUME OUTCOME RELATIONSHIP

3.1 INTRODUCTION

After considering the available evidence on the volume-outcome relationship, it is important to address a series of more methodological questions. These questions include: the completeness and reliability of administrative data, the determination of the volume (surgeon or centre), the statistical methods used, the need for adjustment for case-mix and the causal linkage. This will help not only to place individual studies in perspective, but also to understand why findings may vary across studies.[82]

3.2 RESEARCH QUESTION

The second research question of this project is the following:

What are the different statistical methods used in the literature to assess the volume-outcome relationship in the health care context, and which of these methods specifically assess the question of the causality?

The aim of this chapter is to guide the reader through the different statistical methodologies that have been used so far in that area, to critically assess these methods (pros and cons), and to provide a list of useful methodological references.

3.3 METHODS

3.3.1 Search strategy to select articles

The purpose of this review was to identify methods used to analyze the volume-outcome relationship in the health care context. The limitations of such review need to be acknowledged. While a systematic search of the literature is fully appropriate when answering questions about clinical effects (as in the first question), such search becomes quickly tricky when the aim is to review methods. First because systematic search always rely to some extend to the quality and consistency of the coding of the keywords, and the usual clinical databases (Medline and Embase) are meant for that purpose (keywords not specific enough). Second because good methodological references are often books, not articles, discussing methods that are not specific to the volume outcome relationship but can be used to address it. Therefore, the search was less systematic than in the first chapter, and served mainly to identify and describe the methods used in the volume outcome literature.

Our first step was thus to identify important key words used in analysis of volume-outcome studies. To that end, eight journals with strong methodological emphasis were searched (Statistics in Medicine, Statistical Methods in Medical Research, Biometrics, Journal of Clinical Epidemiology, Journal of Epidemiology and Community Health, Controlled Clinical Trials, Clinical Trials, Journal of biopharmaceutical statistics), with "volume OR outcome" as text in the title, abstract, keywords and text. Based on the articles selected, a list of keywords was identified, and the MESH equivalent of these keywords was searched.

The results are:

- ("Diagnosis-Related Groups"[Mesh] OR "Risk Adjustment"[Mesh])
- "Cluster Analysis"[Mesh]
- "Causality"[Mesh]
- "Longitudinal Studies"[Mesh]

Because not all keywords had a MESH-equivalent, some searches were also based on text only:

- instrumental
- threshold

- selective referral

The last step (and final search algorithm) was to combine each of these MESH terms (or text) to the search strategy from chapter 2 (without the selection of reviews and meta analysis), combined with the "Statistics"[Mesh] term, as papers describing a statistical method are flagged with that term. The whole search strategy is given in Appendix 11 (see Supplement).

From the 1 318 articles identified with that strategy, a first selection was based on the title, to select only those articles pertinent to the volume outcome relationship. A total of 145 articles were selected. After that, all abstracts of those articles were reviewed, and studies lacking the description of the methodology used were excluded. With the review of references and results from searching the grey literature, 10 other studies were added. A total of 98 studies were thus analyzed for their methodology. The graphical flow of selection of articles is presented in Appendix 12.

Two books on the volume outcome relationship were also identified i.e. by Luft et al. and by Seider.[82, 83]

A summary of the sources for the search is presented below:

a. 8 specific methodological journals (using text searches) to obtain key words
b. Medline (using MESH terms and using text fields)
c. The grey literature (Google and Google scholar) for unpublished papers
d. All references from articles selected above
e. Amazon (for books)

3.3.2 Classification of methods

All 98 articles were subsequently classified with regard to two criteria: the design of the study (cross sectional, longitudinal) and the analysis method (regression, choice of covariate, etc.). The oversimplification of such classification scheme compared to a full taxonomy is acknowledged, and should be regarded in light of the aim of the literature search (identify the methods).

3.4 RESULTS: CLASSIFICATION OF METHODS

The classification of methods is based on the exercise undertaken by Luft et al. in 1990, where the authors selected 25 major articles on the volume outcome relationship and classified them into 6 categories, based on the methodology used.[82] The classification by Luft has been extended to include new methods published after 1990 (mainly methods based on econometrics).

The different statistical methods used to analyze the volume-outcome relationship are presented in Appendix 13 (see Supplement), with the number of corresponding articles selected from the search described in section 3.3. It should be noted that these methods are not exclusive: a typical analysis starts with a graphical description of the data, then estimates and tests the relationship in a regression model, and, if a relationship is found, applies specific methods to investigate the causality of the relationship. Again, as mentioned above, articles were categorized based on their main analysis, or on their main methodological emphasis.

Appendix 13 includes:

- The design of the study: the majority of the studies were based on cross sectional designs (93 articles), only five studies were based on longitudinal data
- The statistical methods used:

1. Graphical presentation, funnel plot (1 article)
2. Simple correlation coefficient (1 article)
3. Group by volume and then compare outcomes, with risk adjustments (17 articles)

4. Regression methods
    a. Logistic regression, volume is categorical (42 articles)
    b. Logistic regression, volume is continuous (8 articles)
    c. Cox regression, volume is categorical (9 article)
    d. Poisson regression (2 articles)
    e. Methods specific to hierarchical nature of data (3 articles)
5. Specific to causality: Simultaneous equations modelling (2 articles)
6. Specific to causality: Instrumental variables (5 articles)

3.4.1 Choice of design: cross sectional or longitudinal

The choice of the design has a direct influence on the conclusions that can be drawn from the study. The relation between volume and outcome merely represents an association. Although it is tempting to interpret it as a causal relationship, standard problems in epidemiology arise. For instance, it is not legitimate to say that "as volume increases, mortality falls", and conclude that increasing volume in a hospital will improve outcome. Strictly speaking, if data are taken from a cross section of hospitals observed at a specific time, no conclusions can be drawn on the effect of increasing or decreasing the volume of that hospital on outcome.[82] On the other hand, a design based on the history of mortality at a given hospital, allowing studying the effects of the change of volume on outcome (longitudinal designs) in that hospital, does allow drawing that conclusion.

Nevertheless, it is not surprising that the majority of studies are based on cross sectional data (see Appendix 13). The longitudinal design requires data on the long term (as changes in volume are not expected to occur in a few years time) and also demands more sophisticated statistical analyses.

3.4.2 Description of statistical methods presented in Appendix 13

One interesting article presents a new graphical tool named funnel plot, originally used in quality control process, which allows going beyond the traditional scatter plot (outcome versus volume). This is presented in section 3.5.4.1.

Group by volume and compare outcomes, with risk adjustments

The authors from these articles group hospitals by their patient volume, and then compare outcomes across the different categories. The authors also take into account case mix differences across volume categories through some type of risk adjustment. Usually case mix adjustment is performed via indirect standardization (thus comparing observed versus expected number of outcomes).

The vast majority of the articles used regression methods, which are detailed below:

Logistic regression, volume is categorized

These regression methods involve the use of patient as the unit of observation, although the regression models can also be applied on aggregated data per hospital. Logistic regression was chosen because of the dichotomous nature of the outcome. Volume categories are usually based on the distribution of volume across hospitals (tertiles, quartiles, quintiles) or on other cut off criteria. This is by far the most preferred method of analysis, as shown by the frequencies in Appendix 13.

Logistic regression, volume is continuous

Some authors have performed logistic regression and considered volume as a continuous variable (which it is). Volume or log of volume has been used.

Cox regression (survival analysis)

While short-term outcome (such as in-hospital mortality) can perfectly be handled with logistic regression, long-term outcome (such as 5-year survival) requires more sophisticated techniques from time to event analyses, as some outcome data might be censored. The Cox proportional hazard (PH) model is the most used regression model to deal with censored data.

Specific articles

Some articles were classified apart from the above categories, as they discussed specific methodological questions. One article discussed the correct correlation coefficient for volume outcome relationship studies, and three articles discussed in detail methods to adjust for clustering of data. These include the hierarchical models (also known as multilevel or mixed models) and the Generalized Estimating Equations (GEE) method.

The above articles form what could be called the standard methods for analysing the volume outcome data. Thereupon, other methods, mostly econometrical from origin, were used to specifically test the causality of the relationship. Briefly explained: does the "practice makes perfect hypothesis" (the more one does something, the better one is at doing it) or does the "selective referral hypothesis" (better providers have more volume because they have better quality) explain the relationship? These two hypotheses are discussed in section 3.7.1.

Methods assessing the causality are:

Simultaneous equations modelling

Simultaneous equations modelling is a standard econometrical method used to analyze models with endogenous variable (as volume could be if the selective referral hypothesis is true). There are two equations: one explains outcome as a function of volume, and another which explains volume as a function of outcome. More details can be found in section 3.7.2.

Instrumental variables

This approach is linked to the previous one, as this method specifically constructs new variables, called instruments that are used to adjust the coefficients of the relationship for the possibility of selective referral. A legitimate instrument must be conditionally unrelated with the dependent variable of interest (mortality) except through its effect on the potentially endogenous regressor of interest (volume).[84] These methods are described in section 3.7.3

Longitudinal studies

In these studies, hospitals outcomes data are observed during a certain period of time, assessing what happens in the hospital over time as its volume changes. This has the advantage of keeping other characteristics of hospital and patients unaltered.[33]

This ends the chapter on the classification of methods. The remainder of the chapter discusses in detail the methods presented above, and is organized as follows:

1. presenting the data
2. estimating and testing the relationship (regression models)
3. investigating the causality

3.5 PRESENTING THE DATA

3.5.1 Which outcome?

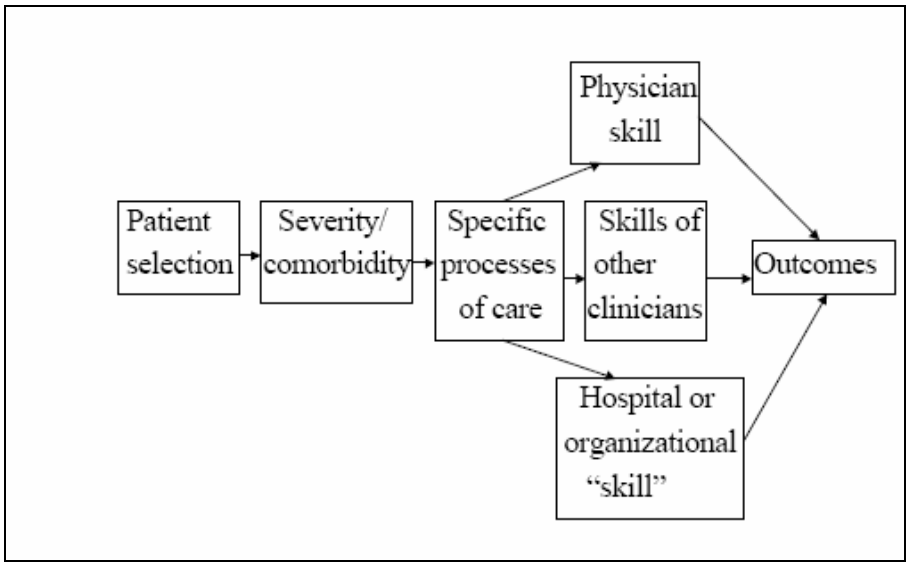
Outcome data can be either continuous (such as length of hospital stay), binary (such as mortality or complication at a certain time point) or as time to event (with censored follow up times). As the vast majority of the literature discusses mortality data at a specific time point, this is also the approach followed in this text, to simplify reading. When the methods differ for continuous or time to event data, this will be mentioned.

3.5.2 Which volume?

While the vast majority of the literature discusses the volume of the hospital, there is a growing body of literature which focuses on the relative importance of the physician volume in contrast, or in addition, to the hospital volume. Importantly, these two volume measures have different implications. While surgeon volume is a good proxy for the surgeon experience (either accumulated or current), there is no equivalent for the hospital volume. Hospital volume might be a proxy for different processes of care, staffing, or organizational structures. If both are available, correlations between hospital volume and physician volume should be investigated in order to produce valid empirical estimates.[82]

The figure below presents the conceptual framework in how the volume of a centre affect the outcome could. It shows that the outcome of a centre is a complex combination of patient selection, patient case mix (severity and co morbidities) and processes of care (choice of treatment, organization). Surgeon volume is just one of the many dimensions of the physician skills. Skills of other physicians than the surgeon can also play a role on the outcome.

Figure 3.1: Conceptual Framework: How Could Volume Affect Quality? In Halm et al. [38]



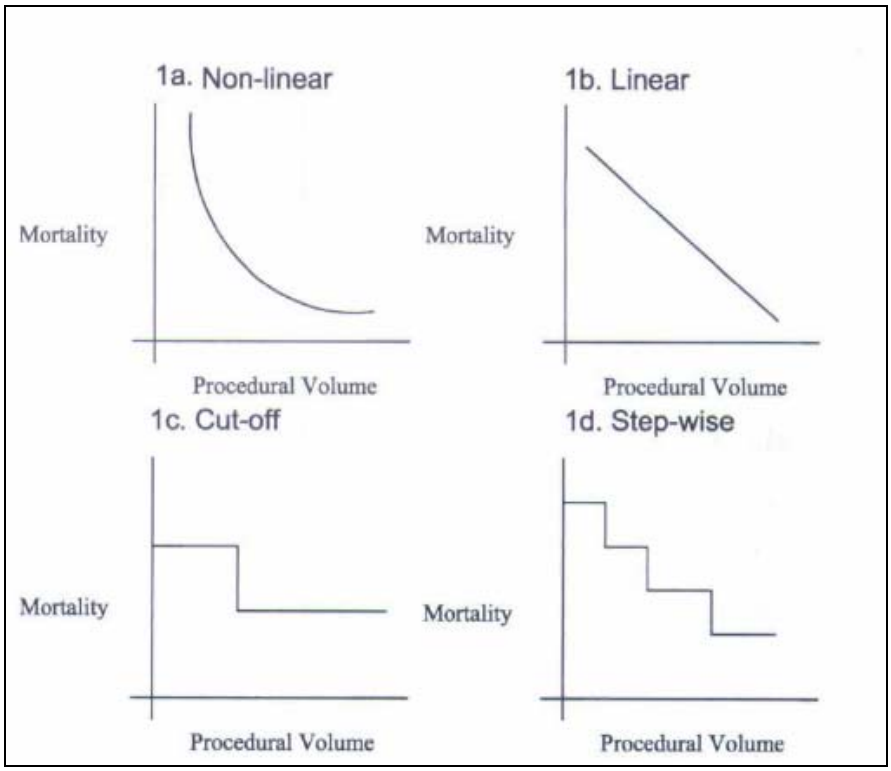
Urbach investigated the relationship between the hospital volume of a specific procedure and the outcome of another procedure, and found some associations.[85]

Some authors also tested the relationship between *previous* surgeon or hospital volume (of previous years), with current outcome, aiming to test the experience at the time of surgery.

3.5.3 The form of the relationship

The nature of the relationship can take many forms, as illustrated in the figure below. It can be either continuous (non linear or linear), have a single cut off (which is the assumption behind selective referral programs) or have multiple steps (stepwise relationship).

Figure 3.2: Possible relationship between volume and outcome, from Christian et al. [30]



3.5.4 Presenting the data

3.5.4.1 Presenting the data graphically

Simple Scatter Plots

The very first obvious step in assessing the volume outcome relationship is to present the data graphically, by the means of a simple scatter plot of all units, having on the axis the volume (or the log of the volume for large variations in volume) and on the y axis an appropriate summary measure of the outcome (proportions of event for binomial data, mean or median for continuous data). This is shown in plot a) in Figure 3.3. To adjust for differences in case-mix between the units, some risk adjusted measures can also be presented (for instance SMR, standardized mortality ratio) (plot c).

Figure 3.3: Examples of different scatter plots, from Betenski et al. [86]

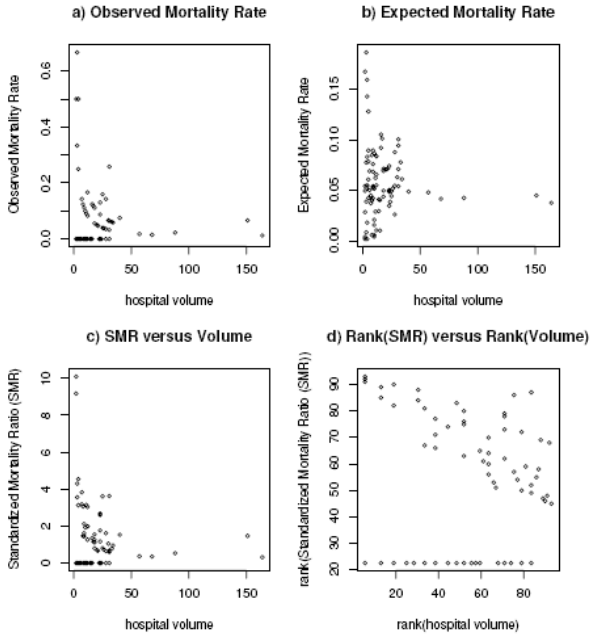


Figure 1. Esophagectomy procedure from 1999 to 2000 UHC data: 93 institutions with 2-year combined volumes ranging from 2 to 164.

The problem with this plot is that the eye is naturally attracted by some of the high mortality rates in the low volume units, and not by the group of low volume units with low mortality rates. Even so, the high mortality rates in the low volume units might as well be within the normal range of the expected variability, based on the small sample size. To get round this issue, another type of plot, named "control charts" (from the statistical process control techniques), has been used in the medical literature. These charts aim to differentiate between « in control » units, showing a common cause of variation, and « out of control » units, exhibiting a special cause of variability, which has then to be investigated further. A specific type of graphical presentation being used in the volume outcome literature is the funnel plot (see Figure 3.2). Others forms of control charts have been presented, but these have limited interest for the volume outcome relationship. [87, 88]

Funnel Plots

Funnel plots have been proposed to assess directly which units are within the expected variability range and which units are not. They have also been suggested as an alternative to overcome statistical problems in the actual ranking of institutions in the league tables published by the NHS.[89-91]

Funnel plots are already a standard tool within meta-analysis, used as graphical check of publication bias. They show the outcome measure plotted against a measure of its precision, so that control limits form a funnel around the target outcome. Figure 3.2 shows the data from the famous "Bristol Inquiry", plotting the observed mortality rate after paediatric surgery against the volume of the institutions, with the superimposition of 95% (~2 SD) and 99.8% (~3 SD) prediction limits around overall mortality rate. This graphic shows that all institutions are within the expected range of variability, except for the data of one centre (Bristol).

Funnel plots have many advantages: the axes are readily interpretable, so that additional information can be added by hand if desired, the eye is naturally drawn to important points that lie outside the funnels, there is no spurious ranking of institutions, and there is clear allowance of additional variability in institutions with small volume.[89] These plots provide thus a good basis for informal graphical check of the volume outcome relationship.

These plots are less helpful when the majority of institutions have a very low volume (one or two interventions) because all units will be concentrated on one dot, and so the size of the dot needs to be adapted to the number of units on that dot. Also, for procedures with low rate of events and units with small sample size, the graphic might give the impression to be "squeezed" if a low volume units has 100% mortality,

The definition of funnel plots has four components. [89] In each unit (hospital or surgeon), r events are observed out of a sample size n (cross sectional binomial data).

1. An indicator (summary statistic) which is the observed proportion of event r/n.
2. A target proportion which is the average event rate θ_0. It is given by the sum of all events divided by the sum of all sample sizes.
3. A measure of the precision, in that case given by the unit sample size n.
4. The control limits that depend of the target θ_0, of the sample size n and of a given p-value. These limits are constructed such that the chance of exceeding these limits for a « in control » unit is p. Usual sets of values for p are p=0,001, p=0,999 corresponding to 3 SD (the usual limits in control charts framework), and p=0,025, p=0,975 corresponding to 2 SD (the usual limits set in the test of hypotheses framework). In the case of binomial cross sectional data, the limits are given by $y_p(\theta_0, n) = \theta_0 + z_p\sqrt{\frac{\theta_0(1-\theta_0)}{n}}$, with z_p as such that $P(Z \leq z_p)=p$ for a standard normal distribution Z ($z_{0.025}$=-1.96).

Figure 3.4: An example of funnel plot, from Spiegelhalter et al. (2005)[89]

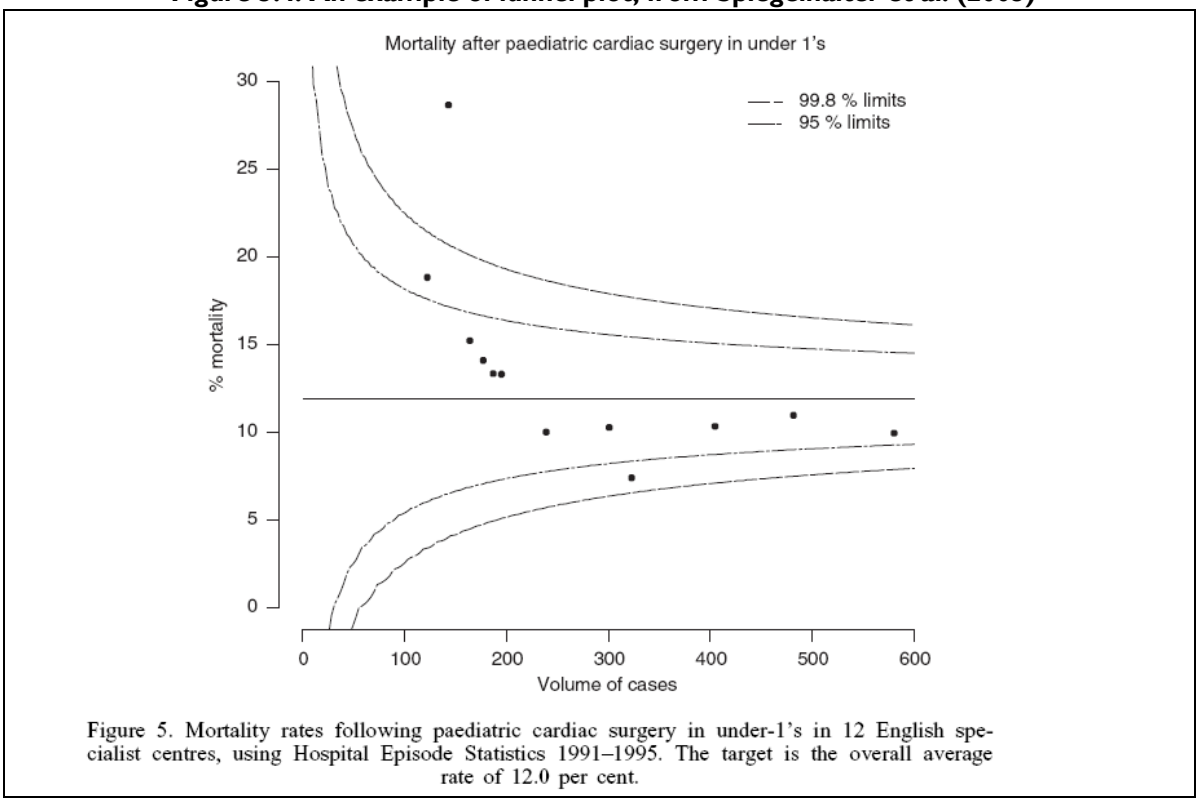


Figure 5. Mortality rates following paediatric cardiac surgery in under-1's in 12 English specialist centres, using Hospital Episode Statistics 1991–1995. The target is the overall average rate of 12.0 per cent.

3.5.4.2 Presenting the data in a table: choice of the volume cut off

Some studies divide the patients into groups of equal volume (for example in tertiles, quartiles or quintiles) and then compare mortality between these groups. The advantage from a statistical perspective is that it ensures that the test comparing the groups has maximal power. But this approach has also drawbacks: [30] there is a great variability in volume between the groups (depending on the distribution of patients within hospitals), potentially ending in meaningless cut off values.

Other studies divide the hospitals into groups of equal volume, based on percentiles or on predefined cut off values (international thresholds). This approach has the advantage of leading to results easier to interpret, but it implicates that low and high volumes should be defined before the analysis. Of course it would be tempting to deliberately select volume threshold after the analysis of data, in order to maximize the difference between the groups. But, this approach is to be avoided, as it makes it impossible to interpret the claimed level of significance (the alpha level of the test).[92] For the same reason, the approach taken by some authors to find the best threshold for volume by testing all different cut off criteria is not valid.[93] It can only serve to generate hypotheses which need to be confirmed on another set of data.

Once data have been presented graphically and summarized in a table format, the next step is to estimate and test the importance of the association.

3.6 ESTIMATING AND TESTING THE VOLUME OUTCOME RELATIONSHIP

Regression models, with outcome as the dependent variable and (a function of) volume as the independent variable, can be used to test the association between the 2 variables. Logistic regression (for binary outcome data) and linear regression (for continuous outcome data) can be used to model individual or aggregated data, the former allowing more flexibility in the choice of the model and in the adjustment for severity of patients. Hierarchical models (synonyms are multilevel model and mixed models) can also take into account the hierarchical nature of health care data (patients nested by physician, and physicians nested by health care provider). These topics are discussed in detail hereafter.

3.6.1.1 Simple regression models

A formal test of the association between volume and outcome can simply be obtained by a logistic regression of the outcome for a patient in a centre on the volume or log volume of that centre.

These models are referred to as conventional logistic regression models, meaning that they do not really account for the hierarchical structure in the data.[94] They easily allow for adjustment of other patient covariates (see section case mix adjustment 3.6.1.2).

When the log of volume is used as explicative factor, the coefficient β has the following attractive interpretation, based on a *relative* change in volume: a small percentage rise x% in sample size leads approximately to a β times x percent change in the odds of death.

When the absolute value of volume is used, the interpretation of the odds ratio (OR) is based on the *absolute* change in volume unit. For each additional unit of volume, the estimated risk for each patient (expressed as the odds of event) is reduced by -100 (1- OR) %. 95% CI and p-values can be derived from all models.

3.6.1.2 The need to adjust for patient severity (Case Mix)

A serious problem in the analysis of the volume outcome relationship is the potential confounding effect of differences in patient severity of illness, as individual factors strongly influence outcomes. The crucial question is whether more (or less) severely ill patients are consistently being treated in high (or low) volume hospitals.

Determining whether the volume outcome relationship is the result of case mix differences would be possible if large numbers of patients were randomly assigned or referred to institutions with varying volume levels.

As this experiment is practically infeasible, the case mix differences must be controlled through various statistical means. Unfortunately, such approaches are never entirely convincing to the sceptical reader. However, a decision must be made whether the likely problems can fully explain the reported results. That is, can plausible biases due to unexplained variables account for the observed results?[82]

There are two general approaches to deal with differences in patient case mix among hospitals. The first approach is to specify a procedure or a diagnosis as carefully and narrowly as possible, resulting in a reasonably homogeneous group of patients. The second approach, which can be combined with the first, is to include variables in the analysis that may capture risk differences among patients included in the study.

The problem of improper case mix adjustment is crucial when the purpose is to compare or to rank institutions (health care profiling). To some extend, the problem of inadequate case mix adjustment is less troublesome if the focus is the pattern of outcome rather than individual hospitals. In order for the case mix to compromise the volume outcome relationship, there must be unmeasured differences in case severity across hospitals, and they must be correlated with volume. Random errors will only reduce one's ability to detect a true relationship. Only if the omitted variables, such as severity, are correlated with explanatory factors, such as volume, will the estimated relationship be biased.[82]

A complete discussion of proper case mix adjustment for health care outcome data is beyond the scope of this chapter. For in-depth discussion we refer to the book by Iezzoni.[95]

Knowing that the method for adjustment and the variables elected as confounders can influence the results, volume effects should always be presented with and without adjustment for case mix, so that the reader can assess the magnitude and direction of adjustment on the estimates.

3.6.1.3 The hierarchical nature of data

Health outcome data have a hierarchical structure by nature. This structure consists of 2 or 3 levels. Figure 3.5 shows the two level structure where several patients are treated by 1 physician or provider. Figure 3.6 illustrates the 3 levels with several patients treated by a physician or provider, and several providers belonging to the same institution. The reality is often more complex, as physicians may perform in different institutions (cross-classified model, Figure 3.5). Although specific models exist to analyze cross-classified structures, their complexity and their lack of availability in standard software prevented them to be used in volume outcome studies. Instead, physicians performing in different institutions are usually considered in the analysis as different physicians at each hospital.[94]

Figure 3.5: An example of 2 level structure (level 1 patient, level 2 provider or institution)

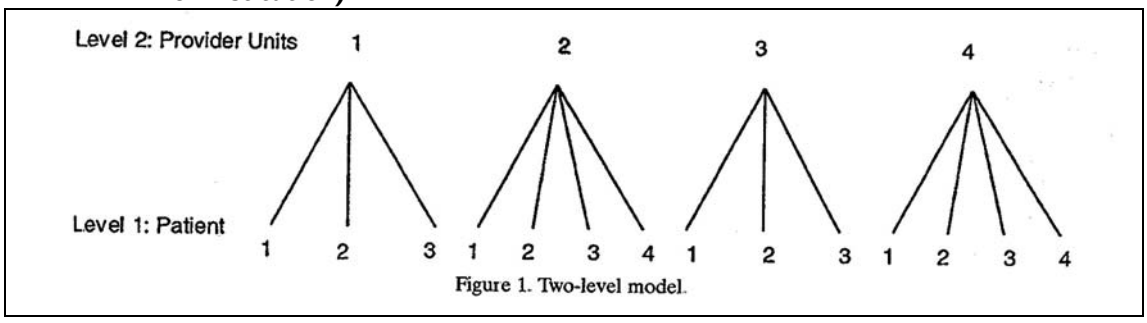


Figure 1. Two-level model.

Figure 3.6: An example of 3 level structure (level 1 patient, level 2 provider, level 3 institution)

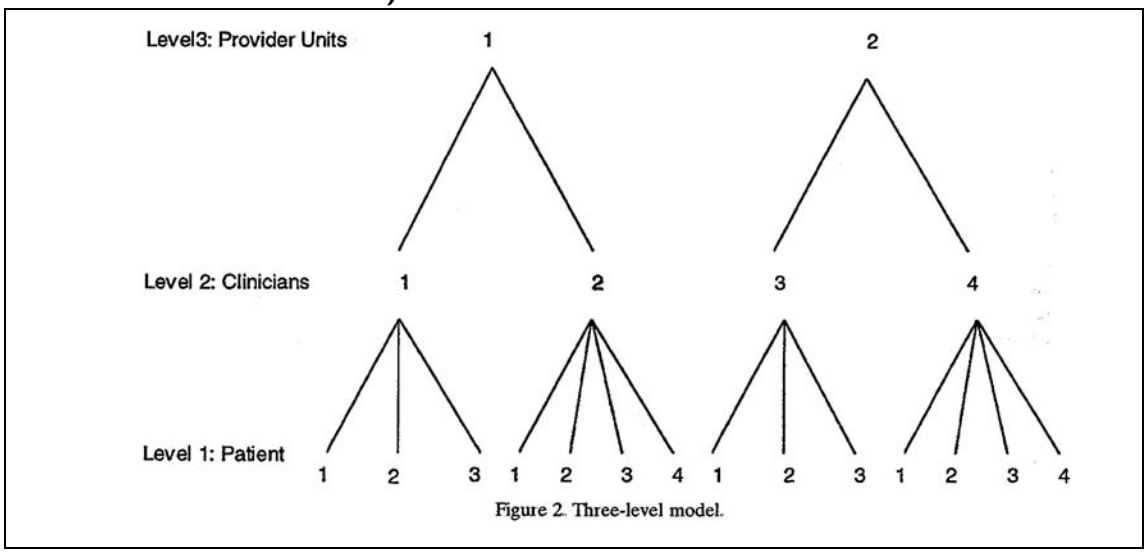


Figure 2. Three-level model.

Figure 3.7: An example of a cross classified 3 level structure (level 1 patient, level 2 provider, level 3 institution)

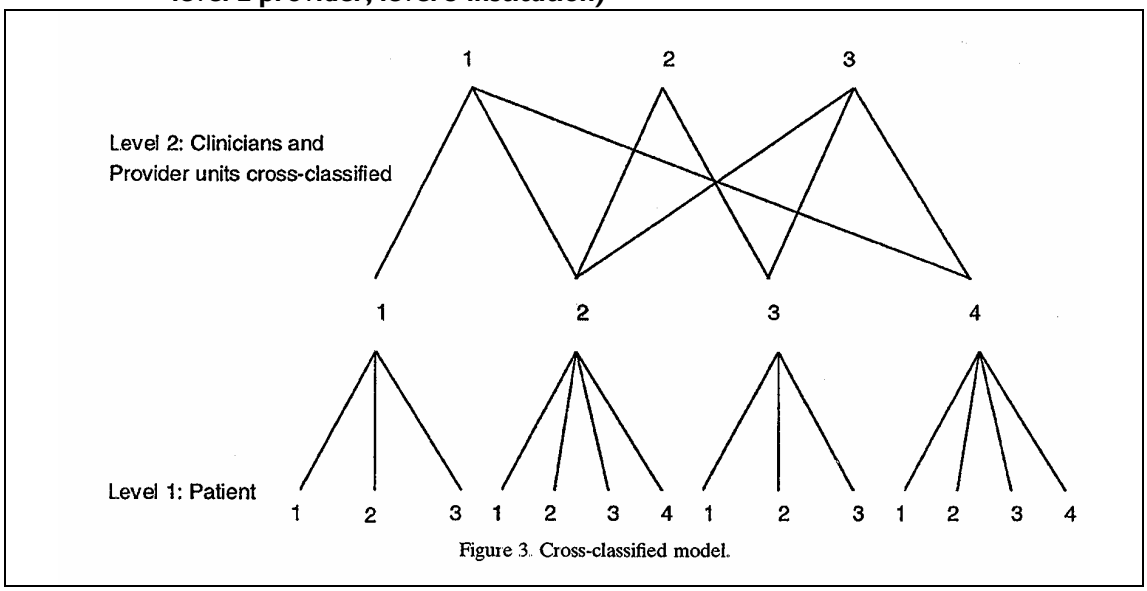


Figure 3. Cross-classified model.

In a typical volume-outcome analysis, the outcome is measured at the level of the individual patient. To account for differences between patients which also influence the outcome of interest, the variation in these factors must be taken into account in the analysis. The usual method is logistic regression analysis, with institutions characteristics attributed to individual patients, and inference of institutions is estimated in a single level multivariable regression model taking other potential confounders into account.

A well known and well described problem with this approach is that it ignores the clustering of patients within hospitals. One of the hypotheses of conventional regression is that observations are independent of each other. This can be violated when data are clustered, because they share other characteristics than the volume and therefore the amount of information present in the data is less than in independent data. The consequence of using conventional logistic regression that does not take into account the clustering of data is that it tends to underestimate the standard error of the regression coefficient, and therefore overestimates statistical significance of apparent effects (standard errors are too small). In other words, an apparent statistically significant relationship using a conventional model might turn out to be non significant when clustering of data is accounted for. This has been extensively described by Urbach and Panageas among many others.[94, 96, 97]

There are two well described methods to account for the effect of clustering of outcome: the random effect (RE) model and the generalized estimating equations (GEE) method.

In the GEE method, the marginal distribution of the outcome is specified, and the dependence between observations is treated as a nuisance parameter (i.e. is not estimated). A correlation matrix that describes the nature of the association within clusters (physicians or centres) is assumed. The resulting robust variance estimates are corrected for intracluster correlation. Although a variety of choice is theoretically appropriate for the correlation matrix, two are most commonly used: the independence correlation matrix and the exchangeable correlation matrix. The independence correlation matrix assumes that outcomes of two patients within the same cluster are independent; whereas the exchangeable correlation structure assumes all pair wise correlations within a cluster are equal. This assumption might be more realistic than the independence assumption. [97]

The RE model (also known as hierarchical model, or multilevel model, or mixed models) has the same form as the conventional model, but it includes a random intercept term that models the hospital specific effect and the physician specific effects. The usual assumption is that hospital specific random effects and physician specific effects are normally distributed and independent from each other. This explicitly models the heterogeneity in outcome across providers and across physicians, which can be quantified. The random effect in the model induces correlations between patients treated by the same physician, and between physicians treating in the same hospital, so that the amount of information is properly estimated and tests of hypotheses have the correct alpha level.

Although both of these methods aim at correctly adjusting for clustering of data, they are not identical and they may sometimes lead to different results, as shown in a comparative study by Panageas.[96]

3.7 THE CAUSALITY ISSUE

3.7.1 Practice makes perfect or selective referral effect?

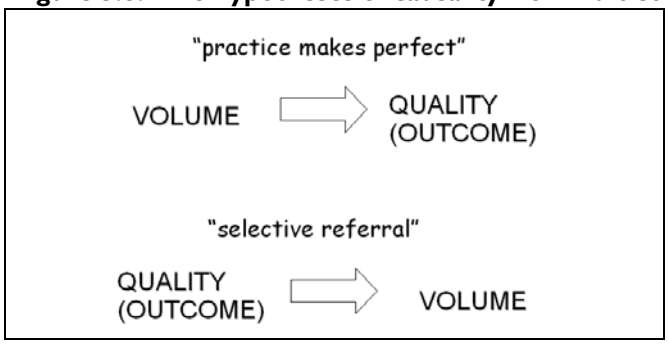
While the vast majority of the volume outcome literature focuses on testing the hypothesis and estimating the volume effect, for policy makers the real question lies elsewhere. Policy decisions in relation with minimal thresholds or referral procedures can only be made after investigation of the causality. Two hypotheses with opposite causal implications have been offered as explanations of the underlying reason for the relationship.

The first causal interpretation, proposed by Luft in 1979 in what is considered the foundation paper of the volume outcome relationship, is the « learning by doing » or « practice makes perfect » explanation.[98] The idea is simple: the more one performs a given task, the better one is at that task. In the health care context, high volume providers perform more interventions, increase their skills and therefore improve outcome. There is certainly a logical argument in favour of that interpretation. In the industry context there is evidence of a learning curve whereby production becomes more efficient with greater experience. Similarly, one would expect a surgical team that performs one open heart procedure a day to be more proficient at it that than a team that performs one a month.[98]

Some years after, in 1987, Luft et al proposed an alternative explanation that they termed « selective referral ».[99] This hypothesis postulates that high quality providers have, on average, a higher volume, simply because patients, perhaps with the advice of a physician, prefer to be admitted to high quality hospitals. The direction of causality is then inversed: it is the high quality that causes the high volume. Even when outcome results are not publicly available, it is plausible that a referral system could function, based on the reputation as « the best in the area » for a particular type of intervention. Alternatively, some providers may develop poor reputation among referring physicians and therefore loose referrals. To quote the analogy of the restaurant given by Luft: a stranger arrives in a town and observes two restaurants with similar prices and location; he expects the one with more customers to be the one with higher quality.[99]

However, it is not the volume of the restaurant that creates the higher quality, but the higher quality that attracts higher volume of customers.

Figure 3.8: Two hypotheses of causality from Luft et al. [99]



Luft et al. explored the plausibility of each hypothesis on a series of diagnoses and procedures, and concluded that both explanations were valid, and that the relative importance of the practice or referral explanation varies by diagnosis or procedure.[99]

The policy implications of the two competing hypotheses are also very different. If the observed pattern reflects only the « practice makes perfect » phenomenon, concentrating patients in selected hospitals will improve outcomes. On the other hand, if the observed pattern is entirely due to selective referral, the concentration of patients is not necessary. As shown by Luft et al, the reality is often less black-and-white, with hypotheses not being mutually exclusive. Given the substantial differences in policy implications, it is important that the analyses try to distinguish both hypotheses.

Luft et al proposed two approaches to investigate which explanation of the two is the more plausible.[99] The first approach is mainly descriptive, and the second uses simultaneous equations modelling.

The first simple approach to explore the relation between volume and outcome is to categorize hospitals by the number of patients in a particular diagnosis or procedure category, and then to examine patterns of selected variables across volume and types of patients. Luft et al proposed three indicators:

1. Transfer into the hospital. The proportion of patients transferred into one hospital from another acute hospital is a direct measure of the selective referral. Luft et al showed that for some procedures, there are marked differences in transfer rates with respect to volume. The authors differentiate the patterns as « strong increasing pattern », « weak increasing pattern », « U shaped », « roughly flat » and « L-shaped ». The strong increasing trends are consistent with the selective referral and inconsistent with the position that practices makes perfect as the only explanation of the volume outcome relation.
2. Transfer to another hospital: the proportion of patients discharged to another hospital rather than to a convalescent facility or home is the flip side of the in transfer rate. Luft shows that for most diagnoses and procedures the transfer rate falls with volume.
3. The risk pattern of patients. This is the final piece of evidence with respect to selective referral. An expected mortality rate based on patients' characteristics can be computed for each hospital.

3.7.2 Use of simultaneous equation modelling

The objective of the second method proposed by Luft is to demonstrate that both hypotheses are viable - and perhaps simultaneous- explanations of the volume outcome relation.[99] The method thus models both equations simultaneously, as explained below.

Figure 3.9: Simultaneous equation modelling from Luft et al [99]

A system of two simultaneous equations

(I) $\text{Volume} = \beta_1 \text{ outcome} + \text{ instruments}_1$

(II) $\text{Outcome} = \beta_2 \text{ volume} + \text{instruments}_2$

The dependent variables in both equations (volume and outcome) are referred as endogenous variables (they are simultaneously determined by the model), and the other variables are referred as exogenous variables (determined outside the model). The fact that the volume variable is endogenous variable has its consequences for the estimation of equation (I), leading to biased estimates of the volume effect.

One equation explains outcome as a function of volume and other factors, and a second equation explains volume as a function of outcome and other factors. To be able to estimate the models parameters simultaneously, a set of instrumental variables is needed. These are a set of exogenous variables that can be convincingly excluded from the first equation and another subset of variables that can be excluded from the second equation. In this type of models, a significant negative coefficient of the volume variable in the equation explaining death rate will support the practice makes perfect hypothesis. A significant negative coefficient on the actual minus expected death rate in the volume equation supports the selective referral pattern hypothesis.

Table 3.1: All theoretical possible results from the two simultaneous equations models proposed by Luft et al.

Practice makes perfect: effect of volume on death rate (β_2)	**Selective Referral pattern: effect of death rate on volume (β_1)**		
	negative	not significant	positive
Negative	Both hypotheses	Practice makes perfect	Practice makes perfect, and counterintuitive selective referral
Not significant	Selective referral	No clear relationship	Counterintuitive selective referral
Positive	Selective referral and counterintuitive practice makes perfect	Counterintuitive practice makes perfect	Both hypotheses counterintuitive

3.7.3 Use of instrumental variables

Some authors have used the instrumental variables in a different context than simultaneous equations modelling. Indeed, their interest lied in estimating consistently the effect of volume on outcome in equation (I), taking into account the endogenous character of that variable. An additional variable is needed, called instrumental variable, assumed to be uncorrelated with the errors of the model but correlated with endogenous variable. The real problem is that it is sometimes far from obvious which variables could act as appropriate instruments, and the quality of the method depends crucially of the quality of the instruments. The use of an instrumental variable, a standard method in economics, is rather new in the medical literature. It has been used for instance to assess the effect of different treatments for acute myocardial infarction,[100] in the context of evaluating the mortality as a measure of quality of hospital care in pneumonia diagnoses,[101] or else to assess the impact of health care acquired infections on length of stay.[102]

In the context of the volume outcome relationship, different instruments have been used. The number of hospital beds [99] [103] [104] and the distance between the patient home and the hospital [84] [105] [106]. Different procedures and conditions have been studied: congestive heart failure [84], heart attack [105], Whipple procedure, CABG and abdominal aortic aneurysm repair[106]. Instrumental variables are still an area of research in statistics.

4 METHODS APPLIED FOR THE ANALYSES

4.1 SOURCE AND LINKAGE OF DATABASES

4.1.1 Minimal Clinical Data – Minimal Financial Data (MCD-MFD)

The registration of the Minimal Clinical Data (MCD) is mandatory for every hospital in Belgium since 1991. This means that for each hospitalized patient, information such as birth date, sex, postal code of domicile and other information such as length of hospital stay, hospital ward and bed type occupation, has to be recorded, along with ICD-9-CM-CM encoding of relevant diagnoses as well as diagnostic and therapeutic procedures performed. Diagnostic and procedure codes are collected per attended hospital department. This inevitably results in a possible redundancy for certain stay specific diagnosis codes causing code frequency counts sometimes to exceed stay counts. After stripping of direct patient-identifying information, records have to be sent biannually to the federal Ministry of Health (MoH). Here, all department registrations are concatenated with assignment of the principal diagnosis of the whole stay, determinant for the APR-DRG-grouper software.

Since 1997, the MCD records are afterwards linked to the Minimal Financial Data (MFD), yearly transmitted by the national health insurance companies (Hicks) to the National Institute for Health and Disability Insurance (NIHDI) and assembling the remuneration costs of each hospital stay. MCD-MFD linkage is performed by a legally instituted 'Technical Cell for the processing of hospital data' (TCT) and requires separately sent correspondence tables containing for each identifiable hospital stay an unique patient pseudonym created by two separately executed hashings: the first by the hospital or HIC respectively and the second by an appointed security advisor of the MOH. This procedure is approved by the Belgian Privacy Commission. The linkage process takes about 2 years to completion and full validation. Linkage percentages increased over the years and exceed nowadays 95% overall.[a] This means that the relationship between treated pathology and the costs to the health care system can be studied, at least for classical hospital admissions. It is important to recognize that the MCD-MFD registry is structured as a relational database encircling 10 separate datasets for the MCD registry and 7 for the MFD registry.

The MCD database also contains records of 'one day' admissions (i.e. patients not staying overnight in the hospital) and outpatients' treatments requiring hospital facilities, however without coupling with billing data yet. The latter is planned for data of the year 2006.

The advantage of the coupled MCD-MFD data is that registration is obligatory for all hospitals (MCD) and all national health insurance companies (MFD).

4.1.2 IMA-AIM data

The purpose of the Common Sickness Funds Agency (IMA) is to organise and manage a common interface to the health care use and patient characteristics data that are collected by all seven Belgian Sickness Funds. The IMA database contains four types of data: data about all reimbursed health care use per attestation per patient; demographic data (e.g. date of birth, gender, community, decease date); data on the insurance status; data on professional status. A full description of the layout of the database and available variables can be found in KCE report 30.[107]

For this study, the Common Sickness Funds Agency was asked to provide only the decease date of patients who were selected in the MCD 2004. Although IMA data also contain information on the use of (neo) adjuvant therapy, it was decided that a supplementary analysis of these data would be beyond the scope of this study.[b]

---

a Expressed as the fraction of the number of stays in MFD data as denominator; stay counts in MFD are always less than stay counts in MCD data since the latter cover all hospital stays, whether or not they were at the expense of the NIHDI.

b Adjuvant therapy is a treatment given after the primary treatment to increase the chances of a cure. Neoadjuvant therapy is given before the primary treatment. Both may include chemotherapy, radiation therapy or hormone therapy.

4.1.3 Data provided by the Belgian Cancer Registry (BCR)

The Belgian Cancer Registry (BCR) has a database containing the following information:

- incidence date (date of first diagnosis, date of first microscopic confirmation of malignancy)
- basis for the diagnosis (histopathologic confirmation, diagnosis based on technical procedures, diagnosis based on tumour markers, diagnosis based on clinical examination only, autopsy)
- primary localisation and histology of the tumour (ICD-O-3, reported in ICD-10 code)
- laterality (for paired organs)
- differentiation grade
- staging (TNM classification)
- WHO score at time of diagnosis (a performance status score)
- treatment (date of first treatment; received and planned treatment).

For each cancer patient, these data are registered in a continuous longitudinal way.[108, 109]

Two important issues for the use of the Cancer Registry database are completeness and validity of the data. In its 2008 incidence report, the BCR defines completeness as "the extent to which all incident cancers in the Belgian population are included in the BCR". For the Flemish Region a complete coverage (>95%) was obtained for the incidence year 2000 while the other regions were only considered as nearly complete for the incidence year 2004.

For some cancer types an under registration could be suspected. In case of pancreas cancer, for example, an under registration of new cancer cases is assumed based on a mortality incidence ratio greater than 1.[108] These will be mainly cancers in a very advanced stage, without histological confirmation, diagnosed in a palliative setting or diagnosed in elderly patients not treated in a hospital. The other cancers which are studied in this KCE report (see Chapter 5, page 45) were not specifically mentioned as being under recorded in the BCR.

Validity or accuracy is defined as "the proportion of cases in a dataset with a given characteristic (e.g. localisation of the tumour, TNM stage, age) which truly have the attribute".[108] The percentage of microscopically verified (% MV) tumours is mentioned as a positive indicator of validity. As shown in Table 4.1, this percentage is lower for cancer of the pancreas than for the other cancers which are analyzed in this report. The Belgian Cancer Registry mentioned that microscopic verification percentages were, in general, lower for cancer of the pancreas, liver and hepatic bile duct, central nervous system and meninges as well as for patients 75 years and older. These tumour localisations are sometimes less accessible for biopsy and techniques can be too invasive or distressing for the patient. This is a well known phenomenon in oncology.[58]

Table 4.1: Percentage of microscopically verified tumours, 2004-2005, Belgium (Belgian Cancer Registry)[109]

	Basis for diagnosis		
Localisation	**% MV**	**% not MV**	**% Missing**
C15 Oesophagus	99.1	0.5	0.2
C18 Colon	98.9	0.6	0.4
C19 Rectosigmoid junction	98.8	0.7	0.4
C25 Pancreas	84.1	10.7	5.0
C34 Bronchus and lung	94.9	4.3	0.6
C50 Breast	99.7	0.1	0.1

Another indicator of data quality is the proportion of records with missing values for certain variables. In the 2004 dataset, 100% completeness was obtained for tumour localisation, histology, behaviour, incidence date, sex and age of the patient.

Basis of diagnosis reached 99.3% completeness. Primary tumour localisation was well specified in 99.9% of the cases, and histology in 95.5%. Data on the WHO performance score and treatment of the tumour were missing in respectively 46 and 39% of cases, which makes these variables unreliable. Information on laterality and stage is often not complete either; 34.0% of cases related to pair organs lack information on laterality; 29.9% of records where stageable tumours are concerned are missing information on the stage (CombStage).[108]

Table 4.2 shows the percentage of well specified laterality and availability of stage information for the tumours that will be analyzed in this report. The Clinical stage (cStage) is based on the available information obtained before resection surgery i.e. by physical examination, radiologic examination and endoscopy. Pathologic stage (pStage) adds additional information gained by histopathologic examination of the tumour. The BCR merges both stages for reporting reasons into the Combined Stage (CombStage). During this merge, the pathologic stage prevails over the clinical stage, except when the clinical TNM is stage IV. In this report, only the Combined TNM stage will be used for risk adjustment. Tumour sites with less surgical treatment such as oesophagus, pancreas and lung have, in general, a higher percentage of cStage and a lower percentage of pStage. For tumours located at colon, rectosigmoid junction, bronchus and lung and breast, at least 70% of records contain information on the CombStage is available. For oesophageal and pancreas tumours, on the other hand, the % CombStage is respectively 60.9 and 64.9% which can cause problems when used for risk adjustment.

Table 4.2: Percentage of well specified laterality and availability of stage information, 2004, Belgium (Belgian Cancer Registry)[58]

Localisation	% Laterality	% cStage	% pStage	% CombStage
C15 Oesophagus	NA	54.4	21.9	62.0
C18 Colon	NA	28.3	73.9	79.5
C19 Rectosigmoid junction	NA	39.9	83.0	91.8
C25 Pancreas	NA	52.2	26.8	65.2
C34 Bronchus and lung	78.7	65.6	18.2	69.8
C50 Breast	92.8	59.5	78.4	87.2

NA = not applicable

4.1.4 Authorization from Privacy Commission

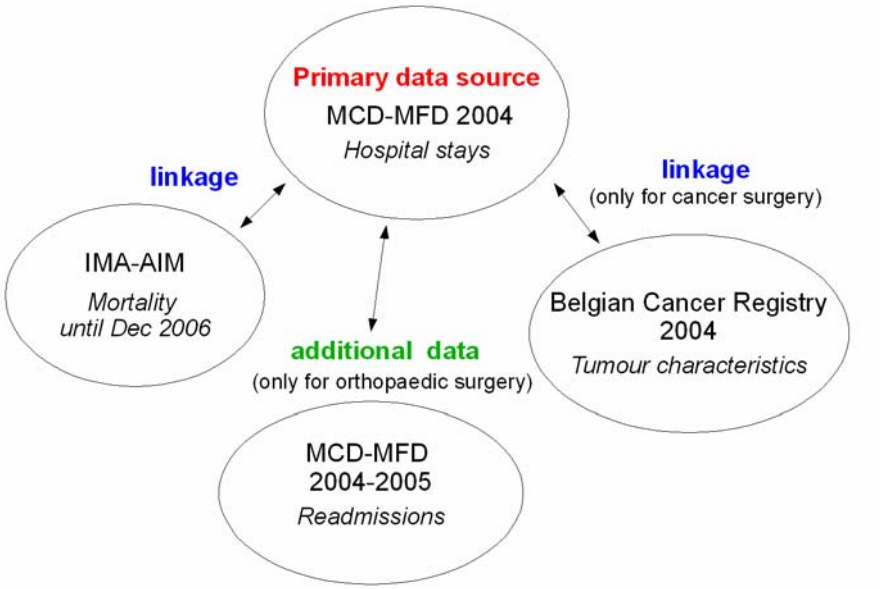
The authorization to access and link these three databases was granted by the Sectorial Committee Social Security and Health of the Belgian Privacy Commission on April 8, 2008.

4.1.5 Linkage of the databases

MCD and MFD were obtained for the year 2004. These were linked to data from the Common Sickness Funds Agency (IMA-AIM) in order to obtain out-of-hospital mortality and socio-demographic characteristics. A second linkage was established with clinical data from the Belgian Cancer Registry (BCR), but only for the oncologic procedures. Readmissions were analysed with MCD-MFD for the year 2005.

Figure 4.1 gives a schematic overview of the linkage between MCD/MFD, IMA-AIM and BCR databases. This linkage was performed in four phases of which detailed flow charts are provided in Appendix 14.

Figure 4.1: Linkage of databases for the KCE study Volume-Outcome



4.2 DATA DEFINITIONS

4.2.1 Primary data selection in MCD 2004

For each condition or procedure, the primary selection of hospitals stays was selected from the MCD-MFD data based on a combination of codes, which are detailed in the results section:

- ICD-9-CM diagnosis code (principal or secondary diagnosis)
- ICD-9-CM procedure code
- NIHDI procedure code.

Principal and secondary diagnosis are defined as follows in the coding handbook of the Belgian Ministry of Health:[110]

- The **principal diagnosis** is defined as the condition which, after examination of the patient, is the principal reason for admission of the patient. The coding handbook specifies that the principal diagnosis is NOT the one present at the time of admission, but rather the one that is made after examining or even operating on the patient.
- The **secondary diagnoses** are defined as the conditions that are present during the hospital stay in addition to the principal diagnosis, or that develop during the stay and have an impact on patient care (e.g. effect on the treatment given or on length of stay) during that stay. Examples of secondary diagnoses are:
  - complications of the principal diagnosis;
  - complications of surgical or medical care during the current stay;
  - associated conditions;
  - active pre-existing conditions.

Only the "classical" hospitalisation stays coupled and validated by the Technical Cell were withheld (no long stay, no one day). All stays fulfilling one of these criteria were included in the primary selection of stays, resulting in a very wide list. This list was afterwards refined in the definition stage.

4.2.2 Definition of procedure

The second step was to define all procedures. To that end, the following descriptive tables were performed, in order to refine the primary selection of data from the previous step:

- details of reasons for inclusion in primary selection
- investigate reasons for differences between inclusion for ICD-9-CM procedure and inclusion for NIHDI procedure (cross tables)
- details of principal and secondary diagnoses
- details of major diagnostic categories (MDC)
- details of APR-DRGs.

As an example, the primary selection of the procedures "carotid endarterectomy (CEA) and carotid stenting (CAS)" was too elaborate (more than 10 000 stays) because it contained all stays with ICD-9-CM procedure code or NIHDI codes, whatever the indication for the procedure. The definition of the procedure selected the stays based on a combination of the principal diagnoses and the selected procedure, and resulted in the inclusion of approximately 3 300 stays.

In addition, some exclusion criteria were applied uniformly across procedures:

- Cancer surgery procedures:
  - Exclude all stays that do not belong to the APR-DRG of interest:
    - Oesophageal cancer surgery: APR-DRG 220 Major stomach, oesophageal and duodenal procedures
    - Peripancreatic cancer surgery: APR-DRG 220 and APR-DRG 260 Pancreas, Liver and Shunt procedures
    - Colon cancer surgery: APR-DRG 221 major small and large bowel procedures
    - Breast cancer surgery APR-DRG 362 mastectomy procedures and APR-DRG 363 breast procedures except mastectomy
    - Lung cancer surgery APR-DRG 120 major respiratory procedures.
- Cardiovascular procedures:
  - Exclude all stays for congenital anomaly. Congenital anomalies are coded with ICD-9-CM diagnosis codes 740-759.
  - Exclude all stays that do not belong to the MDC of the analyzed procedure, i.e. MDC 01 "Diseases and disorders of the nervous system" for CEA/CAS and MDC 05 "Circulatory system" for the other cardiovascular procedures.
- Orthopaedic procedures:
  - Exclude all stays that do not belong to MDC 08 "Diseases and disorders of the musculoskeletal system".

After the definition of each condition or procedure, external data sources were searched for validation of the number of stays. The numbers of procedures were compared, and differences investigated.

4.2.3 Definition of volume

The HOSPITAL volume is the number of stays selected in the definition of the procedure.

The SURGEON volume is the number of stays selected in the definition of the procedure, where the Performer identifier variable (from the MFD) is available in the database. When the same surgeon operates in different hospitals, the surgeon volume is computed across all hospitals (and not the volume in each particular hospital).

If international cut off for volume of centres or surgeon was identified in the literature, these cut off were used in order to facilitate comparison of results. If not cut of volume could be identified, then tertiles (dividing the hospitals in three group of the same size) were used in cardiology and in orthopaedics. In oncology, cut off volume were chosen in order to differentiate the providers (low versus versus medium versus high volume), keeping sufficient number of observations in each group.

The analyses were not intended to determine thresholds but were primarily designed to validate the existence of volume-outcome correlations. Consequently, this precludes making recommendations about specific minimal provider volume thresholds to achieve optimal results.

4.2.4 Definition of outcome

4.2.4.1 Outcome measures

These are the outcome measures which were obtained from the following databases:

1. In-hospital mortality: retrieved from MCD data.
2. Approximate 30-day mortality and 2-year mortality: retrieved from IMA database.
3. Readmission rate: retrieved from MCD data.
   For patients with an index admission in 2004, i.e. who were hospitalized in 2004 to undergo one of the investigated surgical procedures, we disposed of the readmissions until the end of 2005. This information applied to readmissions in both the same hospital as in other hospitals as the index admission.
   It has to be acknowledged that whether the patient is readmitted or not depends on the decision of the physician and is therefore a less reliable outcome to assess the quality of care.
4. Complication rate during index admission or readmission: retrieved from the MCD.

Table 2.1 (page 16) illustrates which outcomes are analyzed for the different procedures.

4.2.4.2 Quality problems of outcome measures

1. As explained in Chapter 2 (see page 15), the approximate 30-day mortality is only an approximation. This means that this outcome measure covers a post-operative window of at least 1 day and at most 60 days.
2. When we compared the (in-hospital) mortality encoded in MCD data with the one registered in IMA data, it was noticed that some patients were encoded as deceased in the MCD data while the IMA data said the contrary. On the whole of the databases, which counted more than 280 000 patients, we counted 1.9% of patients who were deceased in the MCD data 2004 but alive in the IMA data 2004. In these cases, short term mortality was based on the worst case (death in MCD or in IMA database) and long term mortality was based on IMA data only. The rationale behind this decision is that IMA data, which are provided by the Belgian sickness funds, are probably substantially more reliable when the mortality date is concerned. If a patient is encoded as deceased by the sickness funds this means that he no longer benefits any reimbursement of his healthcare costs. An erroneous mortality

date would therefore have enormous implications for the patient. When a hospital, on the other hand, encodes a death in the MCD, this has no consequences for the patient concerned since MCD data are only used for financing the hospital.

3. Complication rate during index admission or readmission: retrieved from the MCD. The problem with this outcome is that it relies heavily on the quality of the hospital coding. Since the recent KCE report by Gillet et al. pointed out that complication data based on MCD cannot be used for hospital benchmarking, only descriptive analyses are performed on the complication rate.[111]

4.3 STATISTICAL ANALYSES ON VOLUME OUTCOME RELATIONSHIP

4.3.1 Funnel Plots

Funnel plots as described in Chapter 2, are used as a first informal assessment of the volume-outcome relationship.

4.3.2 Risk adjustment

For all procedures, different risk factors are taken into account in the regression models (no matter their statistical significance). These factors are

1. patient's age (as a linear variable)
2. patient's gender
3. principal diagnosis of admission (as categories combined based on frequencies and impact on mortality)
4. Charlson score (co-morbidity)
5. stage (for cancer procedures).

The severity of illness (SOI) which is encoded in the MCD was not used for risk adjustment. The reason for this decision is the fact that SOI is encoded at the end of the admission which implies that the patient's complications are taken into account.

4.3.2.1 Charlson score

The Charlson score is a validated score based on patient's co-morbidities, which predicts the 1-year mortality.[112-114] The Charlson score is the sum of some predefined weights attributed to some specific conditions (see Table 2.1). The higher the score, the higher the probability of 1-year mortality is. In the regression models applied in this report, the score is transformed into a five level scale 0, 1, 2, 3 and 4 respectively for scores 0, 1-2, 3-4, 5-6- and > 6, and is analyzed as a continuous variable, as suggested by D'Hoore.[114] The information on co-morbidities is retrieved from the variable 'secondary diagnosis' which is encoded in the MCD.

Table 4.3: Charlson score: Scoring the co-morbidity index from secondary diagnoses

Weight	Conditions	ICD-9-CM code
1	Myocardial infarct	410, 411
	Congestive heart failure	398, 402, 428
	Peripheral vascular disease	440-447
	Dementia	290, 291, 294
	Cerebrovascular disease	430-433, 435
	Chronic pulmonary disease	491-493
	Connective tissue disease	710, 714, 725
	Ulcer disease	531-534
	Mild liver disease	571, 573
2	Hemiplegia	342, 434, 436, 437
	Moderate or severe renal disease	403, 404, 580-586
	Diabetes	250
	Any tumour	140-195
	Leukaemia	204-208
	Lymphoma	200, 202, 203
3	Moderate or severe liver disease	070, 570, 572
6	Metastatic solid tumour	196-199

Source of data: D'hoore [114]

A well known problem of the secondary diagnosis variable in the MCD data is that it is impossible to make the distinction between the complications which occurred during the hospital stay and the co-morbidities that were present at admission, as these are all together encoded in this same variable.[c] For all co-morbidities in the Charlson score, except acute myocardial infarction and cerebrovascular disease, it is a logical assumption that they were present at admission. For acute myocardial infarction and cerebrovascular disease, on the other hand, it is not clear whether they were present at admission or whether they were a complication of cardiovascular procedures. This implies that the Charlson score for patients who underwent a CEA or a PCI can be erroneously high because an AMI which occurred after the procedure was counted as co-morbidity. This problem was partially solved by excluding myocardial infarct and cerebrovascular diseases from the calculation of the Charlson score, which is called the modified Charlson score. This is in line with the approach followed by Birkmeyer in studies on cardiovascular procedures.[115, 116]

When the Charlson score is used for risk adjustment in cancer surgery procedures, it is modified in a different way. In these cases, the two co-morbidity categories specific to cancer, i.e. any tumour and metastatic solid tumour, are excluded from the calculation of the Charlson score, based on methods previously described by others.[117]

4.3.2.2 Cancer stage

The stage of a cancer is a descriptor of how much the cancer has spread. The stage often takes into account the size of a tumour, whether it has invaded adjacent organs, how many lymph nodes it has metastasized to and whether it has spread to distant organs. Cancer stage is important because the stage at diagnosis is a powerful predictor of survival and often determines the kind of treatment.

The TNM system is one of the most commonly used stage systems. It is based on the extent of the tumour (T), spread to the lymph nodes (N), and metastasis (spread to other parts of the body) (M). A number is added to each letter to indicate the size or extent of the tumour and the extent of spread.[118]

[c] The new version of MCD, starting from the second part of 2008 and available from 2010 onward, will contain this information.[111]

TNM combinations correspond to one of five TNM Stages which describe the progression of cancer. In general, the following stages can be distinguished:[118]

- Stage 0: Carcinoma in situ (early cancer that is present only in the layer of cells in which it began).
- Stage I, Stage II, and Stage III: Higher numbers indicate more extensive disease: greater tumour size, and/or spread of the cancer to nearby lymph nodes and/or organs adjacent to the primary tumour.
- Stage IV: The cancer has spread to another organ.

Criteria for stages differ for different types of cancer. For example, oesophagus cancer T2N1M0 is stage IIB;[119] however, colon cancer T4N0M0 is also stage IIB.[120]

4.3.3 Logistic regression models

Logistic regression models are used to assess the effect of the volume of the centre (or of the surgeon) on the outcome (mortality, readmission). The log of the volume (as the natural logarithm LN) is used as explanatory variable in order to study the relative effects. The interpretation of the coefficient is thus a percentage change in the odds of the outcome per percentage increase of the volume. The correlations within the centres (or within the surgeons) are taken into account with Generalized Estimating Equations (GEE), under the assumption of an exchangeable working correlation matrix. All effects of volume are presented with and without adjustment for case mix, to assess the importance of case mix on the relationship, and to enhance transparency of results.

In analyses accounting simultaneously for the volume of surgeons and the volume of hospitals, variances of all estimates are adjusted for correlations of patients within centres (using GEE exchangeable working correlation matrix, as described above). Estimates are not adjusted for clustering of patients within surgeons, because GEE cannot accommodate two levels of hierarchy. Nevertheless, although correlations within hospitals will be therefore somewhat misspecified, estimations are valid because they are based on robust (empirically corrected) standard errors.

4.3.4 Kaplan-Meier survival analysis and Cox PH models

A Kaplan-Meier survival curve is used to analyze the life span of a hip prosthesis. Patients who suffer from a worn out hip prosthesis might require a revision replacement surgery. The survival of the hip prosthesis is measured by identifying the number of days between the first replacement surgery (index admission) that took place in 2004, and the revision replacement surgery. For the latter, we disposed of data until the end of 2005. Data are censored for the following patients: those who did not need a revision by the end of the observation period (31st December 2005); those for whom we did not dispose of a complete follow-up period; and those that died without having had a revision. For total hip replacements (THR), revision rate was calculated at 18 months. For total knee replacement (TKR), this is 12 months. Since information on readmission is only available until the end of 2005, this implies that THR patients who were operated on in the second half of 2004 can not be traced for the entire 18 months. For TKR-patients, on the other hand, we dispose of the entire follow-up time span of 12 months.

For THR, Cox PH models were used to assess the influence of the volume of procedures on the revision rate. Robust sandwich covariance matrix estimates were used to account for the intracluster dependence of observations.

5 RESULTS FOR FIVE CANCER SURGERY PROCEDURES

5.1 THE SELECTED CANCERS

The relationship between volume of interventions and their outcome will be studied for the following procedures:

- oesophageal cancer surgery (including cardia of the stomach) – page 46
- pancreatic cancer surgery – page 65
- colon cancer surgery (including rectosigmoid junction) – page 83
- breast cancer surgery – page 99.
- lung cancer surgery – page 114

This selection is based on the literature review performed in Chapter 2 and on the availability of sufficient cases in one year which implies that rare cancers are not studied.

All analyses were performed on the basis of 2-year survival rate. As there is no censoring of patients during the first two years after surgery, all analyses will be based on binary data (logistic regression). The choice of time horizon was discussed with the external experts who commented on this study. They argued that, for surgeons, short-term mortality is the best outcome for high-risk procedures. Therefore, additional analyses were performed for the two high-risk procedures i.e. oesophageal and pancreatic cancer surgery.

Table 5.1 shows the incidence data of the cancers that will be studied in this chapter as they are reported by the Belgian Cancer Registry.[108] Incidence is expressed as the absolute number of new cases in 2004, and, on the other hand, as the crude (all ages) incidence rate which is the number of new cancer cases observed during a given time period divided by the corresponding number of people in the population at risk, and expressed per 100 000 persons per year.

Table 5.1: Belgian cancer incidence (number of invasive tumours) and crude incidence rate (number of invasive tumours/100 000 person years) (Belgian Cancer Registry, 2004)[58, 109]

Localisation		Cancer incidence (Number)			Crude incidence rate (Number/100 000)	
		Male	Female	Total	Male	Female
C15	Oesophagus	630	234	864	12.4	4.4
C16.0	Cardia	217	59	276	4.3	1.1
C25	Pancreas	583	497	1 080	11.5	9.4
C18	Colon	2 531	2 392	4 923	49.8	45.1
C19	Rectosigmoid junction	318	241	559	6.3	4.5
C34	Bronchus and lung	5 455	1 539	6 994	107.2	29.0
C50	Breast	86	9 369	9 455	1.7	176.5

5.2 OESOPHAGEAL CANCER SURGERY

5.2.1 General description of procedure

As shown in Table 5.1, 1 140 new oesophageal and cardia cancer cases were reported in Belgium in 2004. Surgical removal of the oesophagus (i.e. oesophagectomy) is considered standard treatment for patients with resectable oesophageal cancer. Although recent Belgian guidelines recommend that oesophagectomy should not be performed with palliative intent in patients with oesophageal cancer, it is sometimes a way to relieve symptoms but only in carefully selected patients.[119, 121]

5.2.2 Primary data selection in MCD 2004

First selection of data was based on ICD-9-CM diagnostic and procedure codes as proposed by Dimick et al. and by Christian et al., and on NIHDI procedure codes.[93, 122]

Following the recommendations of the Belgian College of Oncology and the example of others who published on the volume-outcome association in oesophageal cancer surgery, patients with the diagnosis of cardia tumour (i.e. the proximal part of the stomach) were explicitly included.[93, 119, 121, 123, 124]

This selection resulted in a total of 4 565 stays (Table 5.2). This selection is too broad since it also includes patients with a secondary diagnosis of malignant neoplasm of oesophagus. This implies that patients with oesophageal cancer as pre-existing condition are also included (see section 4.2.1, page 39). The definition of the population studied (i.e. patients with a principal diagnosis of malignant neoplasm of oesophagus or cardia, with a surgical resection of the oesophagus) will be defined in the next section.

Table 5.2: Oesophageal cancer surgery: Primary data selection in MCD 2004

SELECTION 1	SELECTION 2	SELECTION 3
ICD-9-CM (principal or secondary) diagnosis code	**ICD-9-CM procedure code**	**NIHDI procedure code**
150x Malignant neoplasm of oesophagus	4240 Oesophagectomy, not otherwise specified	228023 Thoracic or thoracoabdominal oesophagectomy or gastro-oesophagectomy, at one surgery time.
1510 Malignant neoplasm of cardia	4241 Partial oesophagectomy	228185 Subtotal oesophagectomy until the level of the aortic arch, with restoration of continuity.
	4242 Total oesophagectomy	
STAYS SELECTED = 4 476 stays	**STAYS SELECTED = 370 stays**	**STAYS SELECTED = 436 stays**
TOTAL STAYS SELECTED= 4 565 stays (selection 1 OR selection 2 OR selection 3)		

5.2.3 Definition of procedure

5.2.3.1 Primary hospital stays in Minimal Clinical Data

A total of 2 084 of the 4 565 selected hospital stays had a principal diagnosis of malignant neoplasm of oesophagus (74% i.e. 1 545 stays) or cardia (26% i.e. 539 stays). As shown in Table 5.3, 378 stays (18.1% of 2 084 stays) corresponding to 377 patients (26.9% of 1 401 patients) had a procedure for oesophageal cancer surgery as described above (with the addition of the ICD-9-CM procedure 43.99 "Oesophagogastrectomy NOS" which is also to be considered as oesophageal cancer surgery, but was not present in our initial selection).

Table 5.3: Oesophageal cancer surgery: Principal diagnosis and percentage of cancer surgery (per stay and per patient)

Principal diagnosis	Total Number hospital stays	Stays with oesophageal cancer surgery		Total Number patients	Patients with oesophageal cancer surgery	
		Nbr	Pct		Nbr	Pct
150 Malignant neoplasm of oesophagus	1 545	245	15.9	1 005	244	24.3
151 Malignant neoplasm of cardia	539	133	24.7	396	133	33.6
Total	2 084	378	18.1	1 401	377	26.9

The APR-DRGs of these 378 stays are described in Table 5.4.

Although all these 378 stays had a principal diagnosis of malignant neoplasm of oesophagus or cardia and a surgical resection of the oesophagus, they were not all classified in the APR-DRG 220 "Major stomach, oesophageal & duodenal procedures". In some cases, for example, where a tracheostomy was performed, the APR-DRG changed into 004. To create a study population of patients which is as homogenous as possible, it was decided to retain only the 343 stays (corresponding to 342 patients) that are included in the APR-DRG 220.

Table 5.4: Oesophageal cancer surgery: All APR-DRGs of stays with a principal diagnosis of a malignant neoplasm of the oesophagus or cardia AND with surgical resection of the oesophagus

APR-DRG	Number	Percent
004-TRACHEOSTOMY EXCEPT FOR FACE, MOUTH & NECK DIAGNOSES / p3 - P	33	8.73
220-MAJOR STOMACH, ESOPHAGEAL & DUODENAL PROCEDURES / 6 - P	**343**	**90.74**
222-MINOR STOMACH, ESOPHAGEAL & DUODENAL PROCEDURES / 6 – P	1	0.26
226-ANAL & STOMAL PROCEDURES / 6 - P	1	0.26
TOTAL	378	100.00

Results of definition: 343 stays with oesophageal cancer surgery were selected in the Minimal Clinical Data. These stays concerned 342 individual patients and 72 centres.

5.2.3.2 *Linkage with data from IMA and BCR*

The Minimal Clinical Data (MCD) data were linked with the Common Sickness Funds Agency (IMA) database and the Belgian Cancer Registry (BCR) database. Table 5.5 shows that 95.9% of 343 MCD stays could be linked with IMA data and 79.9% with BCR data (on the basis of ICD-10 codes C15 and C16.0). As was to be expected on the basis of the percentage in Table 4.2 (see page 38) information on tumour stage could only be retrieved for 67.4% of 343 MCD stays. However, this percentage varies strongly among the hospitals as is shown in Figure 5.1 where the red columns indicate the procedures for which stage is missing.

Table 5.5: Oesophageal cancer surgery: Percentage of linkage of MCD data with data from IMA and BCR

	Number	Percent
Number of stays in MCD selection	343	100
Linkage with IMA	329	95.9
Linkage with BCR	274	79.9
Linkage with BCR and information on stage	231	67.4

Results of linkage: 274 patients with oesophageal cancer surgery could be linked with Belgian Cancer Registry data. For 231 of these patients, the BCR contained information on the stage of the tumour.

5.2.3.3 *Patient and tumour characteristics*

Tumour location is defined according to the ICD-10 code:

- C15 for malignant neoplasm of oesophagus
- C16.0 for malignant neoplasm of cardia.

Two main types of oesophageal cancer are squamous cell carcinoma and adenocarcinoma. Tumour histology is according the histological groups defined by the International Agency for Research on Cancer (IARC) on the basis of the following ICD-O-3 codes:[125]

- squamous cell carcinoma: 8050-8078, 8083-8084
- adenocarcinoma: 8140-8141, 8143-8145, 8190-8231, 8260-8263, 8310, 8401, 8480-8490, 8550-8551, 8570-8574, 8576.

For a detailed description of TNM classification and TNM stages for oesophageal cancer, we refer to the Belgian national guidelines.[119, 121]

As shown in Table 5.6, the mean age of patients was 64 years (median age is 65 years) and 80.2% of them were men. The majority of patients had an oesophagus cancer (69.3%). The predominant histological type of tumour was adenocarcinoma (67.1%). Twenty-two per cent of patients had stage 1 oesophageal cancer, 32.5% had stage 2, 37.2% had stage 3 and 8.2% had stage 4 disease.

Table 5.6: Oesophageal cancer surgery: Patient and tumour characteristics

		Number	Percent
Age (years)			
	mean/median	63.9/65	
	Std	10.2	
	Total	*343*	
Sex			
	Male	275	80.17
	Female	68	19.83
	Total	343	
Tumour location			
	Oesophagus	190	69.3
	Cardia	84	30.7
	Total	*274*	
Tumour histology			
	Squamous Cell Carcinoma	75	27.4
	Adenocarcinoma	184	67.1
	Other	15	5.5
	Total	*274*	
Tumour stage			
	I	51	22.1
	II	75	32.5
	III	86	37.3
	IV	19	8.2
	Total	*231*	

5.2.4 Definition of volume

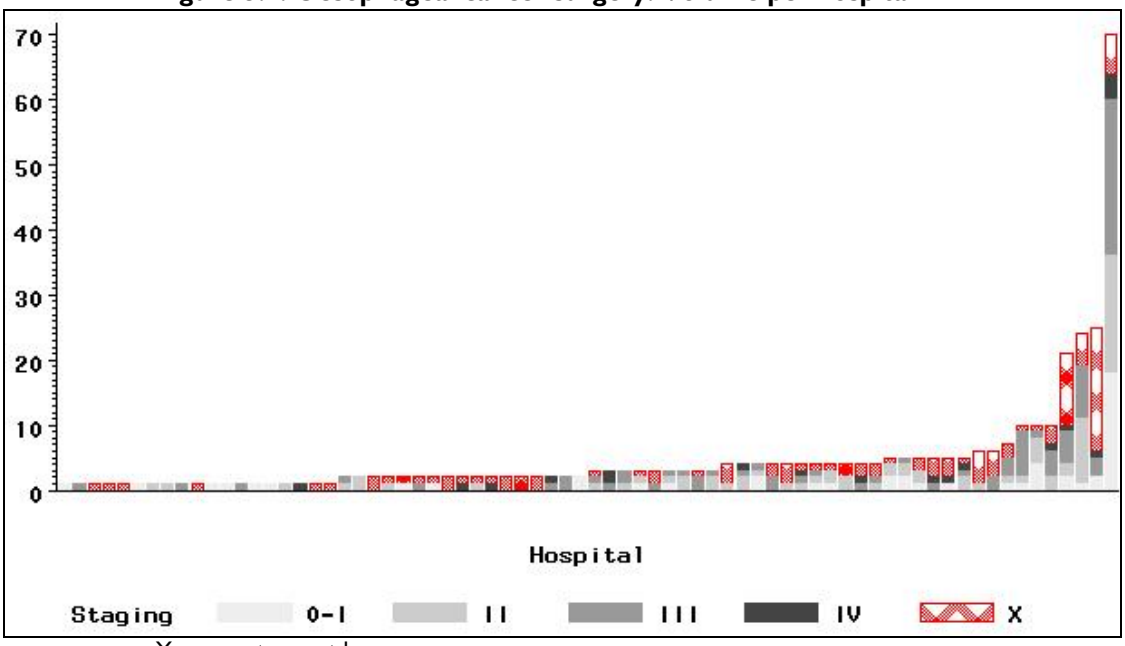
As shown in Table 5.7, these 343 procedures were performed in 72 centres by 99 surgeons (some surgeons operate in more than 1 hospital). Three quarters of the hospitals performed 4 procedures or less. Four hospitals and six surgeons had an annual volume higher than 10 procedures (see Figure 5.1 and Figure 5.2).

Table 5.7: Oesophageal cancer surgery: Summary measures of volume per hospital and per surgeon

	Number	mean	min	25th Pctl	50th Pctl	75th Pctl	Max
Hospitals	72	4.8	1.0	1.0	2.5	4.0	70.0
Surgeons*	99	3.1	1.0	1.0	1.0	3.0	41.0

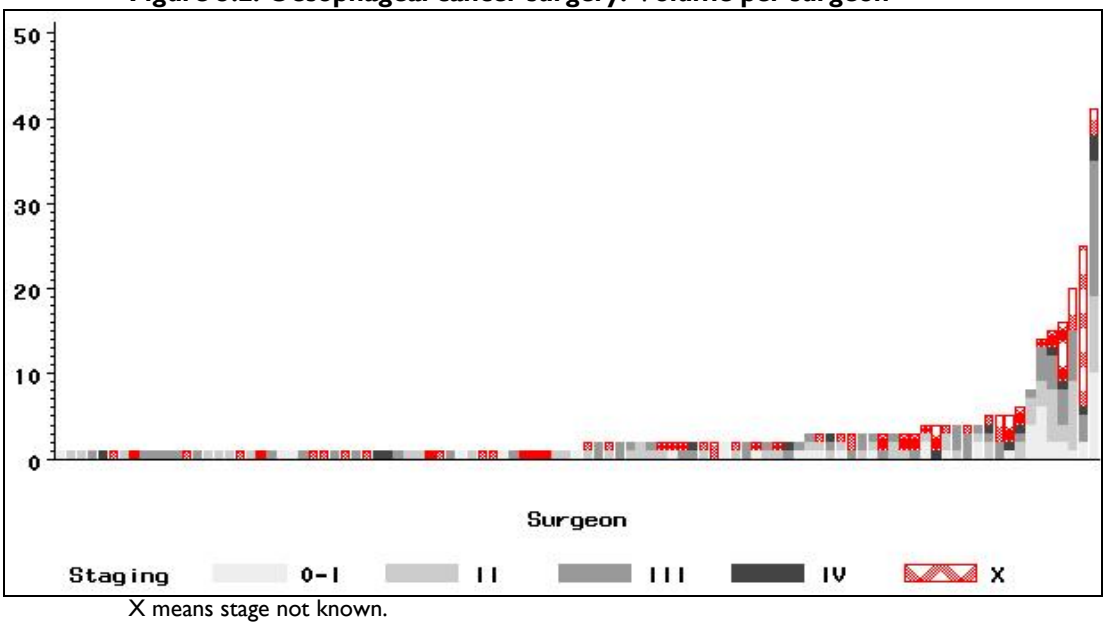
* For 34 procedures (9.9%) the information on surgeon ID was missing.

Figure 5.1: Oesophageal cancer surgery: Volume per hospital



X means stage not known.

Figure 5.2: Oesophageal cancer surgery: Volume per surgeon



X means stage not known.

5.2.5 Definition of outcomes

5.2.5.1 Mortality

Because of errors in the MCD (discussed in section 4.2.4.2 page 41), in-hospital mortality was not congruous with approximate 30-day mortality. We therefore merged both sources of information into one outcome measure i.e. "in-hospital and approximate 30-day mortality". Note that the 30-day mortality is an approximation i.e. minimum 1 day and maximum 60 days, as explained in Chapter 2 (see section 2.3.2.2, page 15).

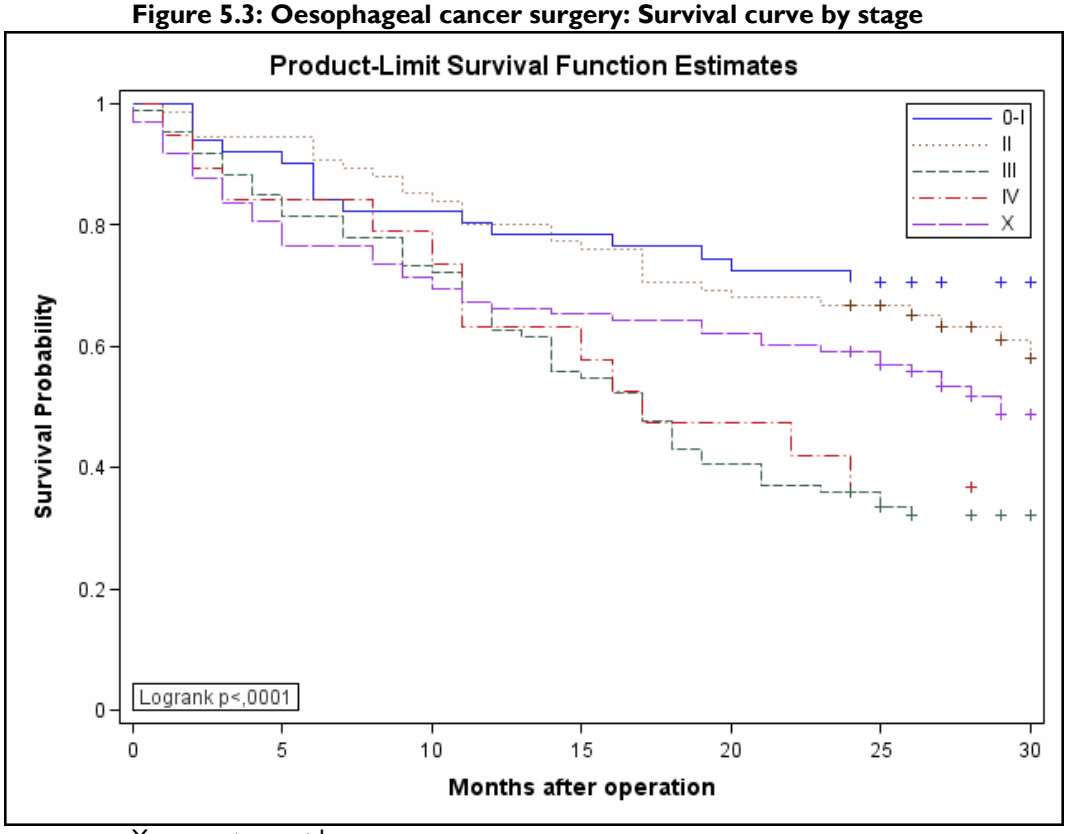
Table 5.8 shows that in-hospital and approximate 30-day mortality was 9.1% after oesophagectomy for cancer. Mortality at two years was 44.7%. Median survival time was 29 months (not shown in Table 5.8).

Table 5.8: Oesophageal cancer surgery: Mortality results

	Number	Number deaths	Percent (%) deaths
In-hospital and approximate 30-day mortality	329	30	9.1
3 months mortality	329	37	11.2
1-year mortality	329	98	29.8
2-year mortality	329	147	44.7

Figure 5.3 shows clearly how survival improves with decreasing disease stage. In line with expectations, patients with stage 3 and 4 disease have the worst survival curves while stage 1 and 2 patients have the best survival. Patients with unknown disease stage (in the BCR data), i.e. the purple line in Figure 5.3, are in between these two groups.

Figure 5.3: Oesophageal cancer surgery: Survival curve by stage



X means stage not known.

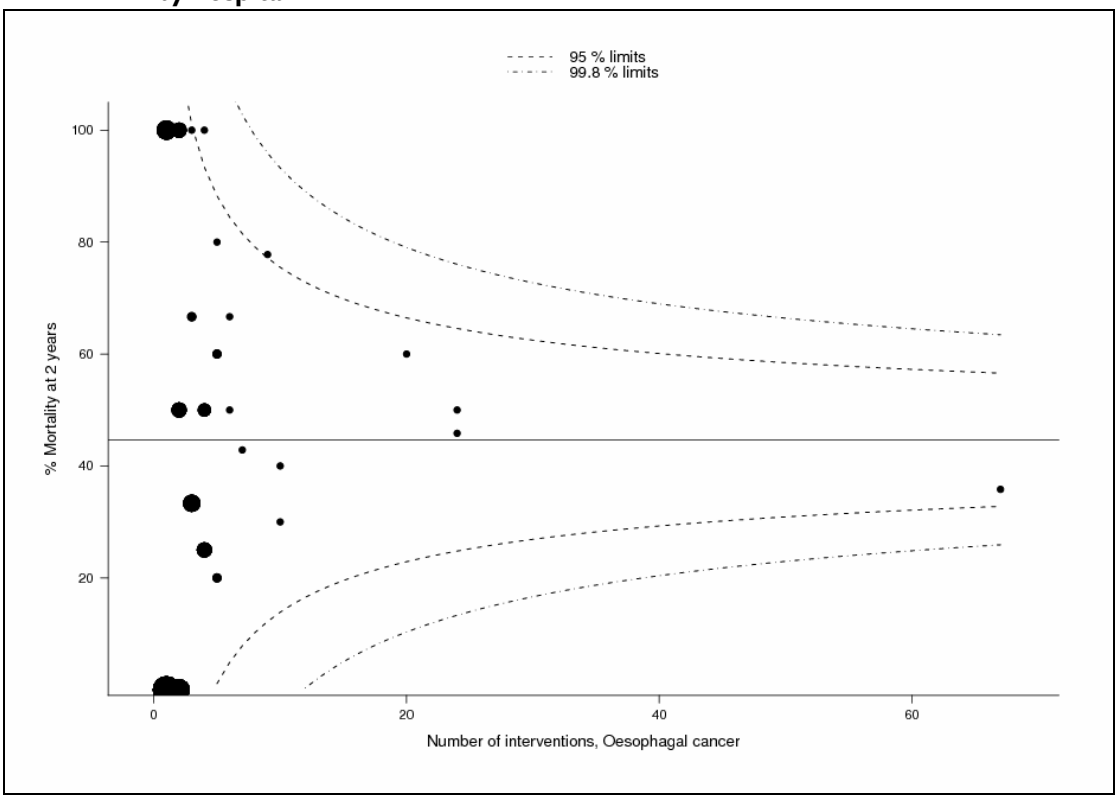
5.2.6 Volume outcome relationship

5.2.6.1 Analysis by hospital

MORTALITY RATE, BY HOSPITAL

Figure 5.4 presents the funnel plot of the number of oesphagectomies and the 2-year mortality (observed, i.e. without risk adjustment). The size of the point is proportional to the number of hospitals with the same volume and the same outcome. The horizontal line represents the overall 2-year mortality i.e. 44.7%. None of the hospitals are outside the 99.8% limits of variability.

Figure 5.4: Oesophageal cancer surgery: Funnel plot of 2-year mortality rate, by hospital



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

Table 5.9 presents the descriptive differences in case-mix in function of the volume of oesophageal cancer procedures per hospital. The 70 centres have been divided into five groups according to their annual volume of procedures. The cut off limits are somewhat arbitrary, as they are not meant to divide the hospitals into five groups of equal size, but to summarize the information and to differentiate the groups as well as possible. 2-year mortality was 36% in lowest volume hospitals (1-2 procedures per year) and more than 40% in all other volume categories. There is no evidence that case mix is related to volume since patients having surgery at high- and low-volume hospitals were very similar with respect to age, gender and co morbidities (i.e. Charlson score). The distribution of the tumour stage among the hospital categories is very heterogeneous and difficult to summarize.

Table 5.9: Oesophageal cancer surgery: Differences in case-mix and outcomes by hospital volume

		Hospital volume					
		1-2/yr	3-4/yr	5-9/yr	10-19/yr	≥20/yr	All
Number hospitals		34	20	9	3	4	70*
Number stays		50	67	48	29**	135	329
Gender	%male	72.0	83.6	83.3	65.5	82.2	79.6
Age (years)	Mean	65.8	64.2	65.9	64.2	62.3	63.9
	Std	8.9	10.8	9.5	10.8	10.5	10.3
Charlson ≥ 3	%	12.0	20.9	8.3	13.8	9.6	12.5
Charlson	Mean	0.9	1.3	1.0	0.8	0.7	0.9
	Std	1.3	1.5	1.1	1.4	1.2	1.3
Severity index (APR-DRG)							
2	%	6.0	.	6.3	3.4	27.4	13.4
3	%	54.0	47.8	39.6	48.3	40.0	44.4
4	%	40.0	52.2	54.2	48.3	32.6	42.2
Principal Diagnosis (MCD)							
Oesophagus tumour	%	44.0	62.7	56.3	55.2	75.6	63.5
Cardia tumour	%	56.0	37.3	43.8	44.8	24.4	36.5
Length of stay	Median	22.0	23.0	24.5	19.0	19.0	21.0
In-hospital + approximate 30-day mortality	%	8.0	14.9	12.5	3.4	6.7	9.1
2-year mortality	**%**	**36.0**	**47.8**	**50.0**	**48.3**	**43.7**	**44.7**
Missing data BCR (stage)	%	40.0	26.9	41.7	13.8	26.7	29.8
Stage							
0-I	%	33.3	12.2	25.0	20.0	23.2	22.1
II	%	26.7	40.8	35.7	28.0	30.3	32.5
III	%	26.7	36.7	28.6	48.0	40.4	37.2
IV	%	13.3	10.2	10.7	4.0	6.1	8.2
Histology							
Adenocarcinoma	%	77.5	70.4	71.1	50.0	64.7	67.2
Other	%	5.0	1.9	7.9	11.5	5.2	5.5
Squamous Cell Carcinoma	%	17.5	27.8	21.1	38.5	30.2	27.4

* Only 70 centres (and not 72) because 2 centres had no data linked with IMA data

** Only 29 (and not 30) because volume based on all stays (and for one centre 1 stays was not linked with IMA data)

RELATION BETWEEN HOSPITAL VOLUME AND 2-YEAR MORTALITY RATE

Analyses were performed with international thresholds i.e. 6 oesophagectomies/year issued by the US Agency for Healthcare Research and Quality (AHRQ). Descriptive results using the 13/year threshold issued by the US Leapfrog Group are also presented.[29, 126] Table 5.11 contains results from a logistic regression model (with GEE adjustment for clustering of patients within hospitals) used to examine the relation between hospital volume and 2-year mortality following oesophageal cancer surgery.

The first model was performed without adjustment for case mix. These results are equivalent to the percentages presented in the previous table: the data suggest that there is no difference between the two groups 0.87 95%CI (0.55, 1.38)

In the second model, an adjustment was made for patient characteristics which were retrieved from the MCD data i.e. sex, age, principal diagnosis and Charlson score. This adjustment barely changes the results of the volume effect on 2-year mortality

In addition, the third model also adjusted for tumour characteristics such as stage and histology of the oesophageal cancer. Again this adjustment does not change the impact of the volume category on the 2-year mortality: odds ratio and 95% CI 0.89 (0.57, 1.40). In this model, patients with a higher Charlson score have a higher mortality rate (OR = 1.44, 95%CI 1.10-1.88). Disease stage also proves to be an important predictor of 2-year mortality. Patients with stage 3 and stage 4 disease have significant higher 2-year mortality than those with stage 1 disease. Tumour histology, however, is not a statically significant predictor of mortality.

Table 5.10: Oesophageal cancer surgery: 2-year mortality per hospital based on international volume thresholds

Cut off		Hospital volume category	Number centres	2-year mortality		
				Number cases	Number deaths	%
AHRQ cut off	6/year	1-5/year	60	146	64	43.8
		≥ 6/year	10	183	83	45.4
Leapfrog cut off	13/year	1-12/year	66	194	88	45.4
		≥ 13/year	4	135	59	43.7

Table 5.11: Oesophageal cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality

Factor		OR	95%CI	
Model without adjustment for case mix				
Hospital Volume	1-5/year vs ≥ 6/year	0.87	0.55	1.38
Model with adjustment for patient characteristics (MCD data)				
Hospital Volume	1-5/year vs ≥ 6/year	0.82	0.53	1.26
Sex	male vs female	1.21	0.63	2.34
Age	increase of 1 yr	1.01	0.99	1.03
Principal Diagnosis	Oesophagus vs cardia tumour	0.83	0.55	1.27
Charlson score	increase of 1	1.29	1.01	1.65
Model with adjustment for patient and tumour characteristics (MCD and BCR data)				
Hospital Volume	1-5/year vs ≥ 6/year	0.89	0.57	1.40
Sex	male vs female	1.42	0.67	3.00
Age	increase of 1 yr	1.02	1.00	1.04
Principal Diagnosis	Oesophagus vs cardia tumour	0.67	0.41	1.08
Charlson score	increase of 1	1.45	1.12	1.88
Stage	unknown	2.18	1.03	4.63
	IV	4.04	1.30	12.57
	III	4.77	2.44	9.35
	II	1.20	0.54	2.65
	0-I	1.00	1.00	1.00
Histology	unknown	0.40	0.14	1.19
	Adenocarcinoma	0.62	0.36	1.04
	Other	1.53	0.48	4.93
	Squamous Cell	1.00	1.00	1.00

Sensitivity analyses with the log of the volume as a continuous variable in the model confirmed these results (no linear effect of log of volume on mortality, data not shown).

5.2.6.2 Analysis by surgeon

MORTALITY RATE, BY SURGEON

The funnel plot of the relationship between volume of surgeon and respectively 3 months and 2-year mortality are presented in Figure 5.5 and in Figure 5.6.

Figure 5.5: Oesophageal cancer surgery: Funnel plot of in-hospital and approximate 90-day mortality rate, by surgeon

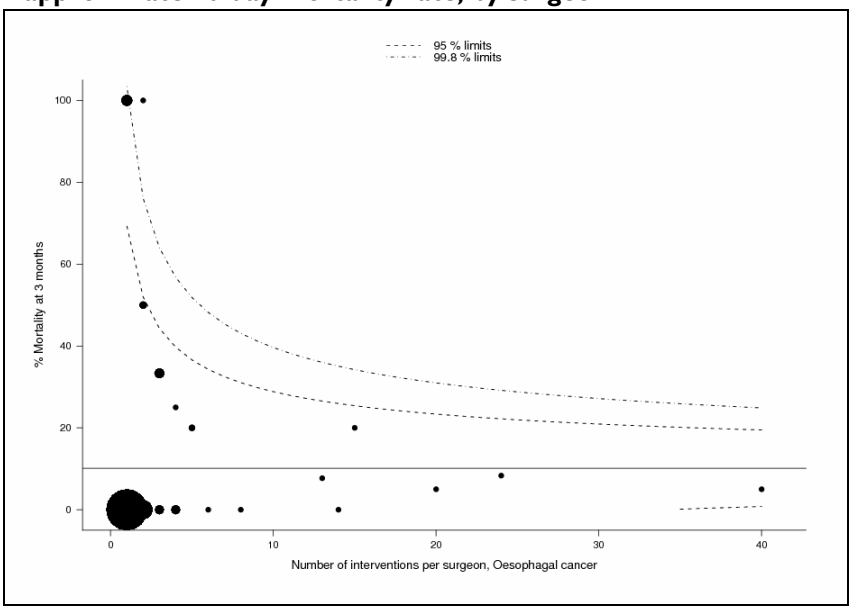
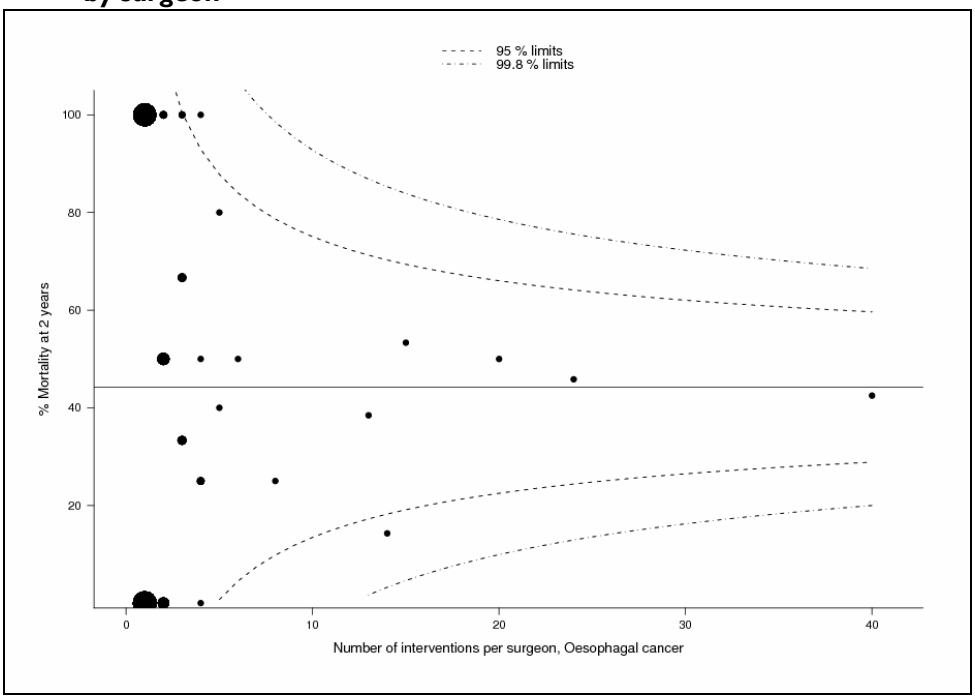


Figure 5.6: Oesophageal cancer surgery: Funnel plot of 2-year mortality rate, by surgeon



RELATION BETWEEN SURGEON VOLUME AND 2-YEAR MORTALITY RATE

When international hospital volume thresholds (AHRQ or Leapfrog) are applied on surgeons, low volume surgeons (1-5/year) have 13.5% 3 month mortality rate and a 46.8% 2-year mortality rate while high volume (≥ 6/year) surgeons have a 6.4% 3 month mortality and a 41.3% mortality rate (Table 5.12). Survival curve is presented in

Figure 5.7. Percentage of are also presented based on the leapfrog cut off (13/year)

Table 5.12: Oesophageal cancer surgery: mortality per surgeon based on international volume thresholds

International cut off	Surgeon volume category	Number surgeons	Number cases	3-months mortality		2-year mortality	
				Number deaths	%	Number deaths	%
AHRQ cut off 6/year	1-5/year	88	156	21	13.5	73	46.8
	≥ 6/year	8	140	9	6.4	58	41.3
Leapfrog cut off 13/year	1-12/year	90	170	21	12.4	78	45.9
	≥ 13/year	6	126	9	7.1	53	42.1

Figure 5.7: Oesophageal cancer surgery: Survival curve by volume of surgeons (based on AHRQ cut off 6/year

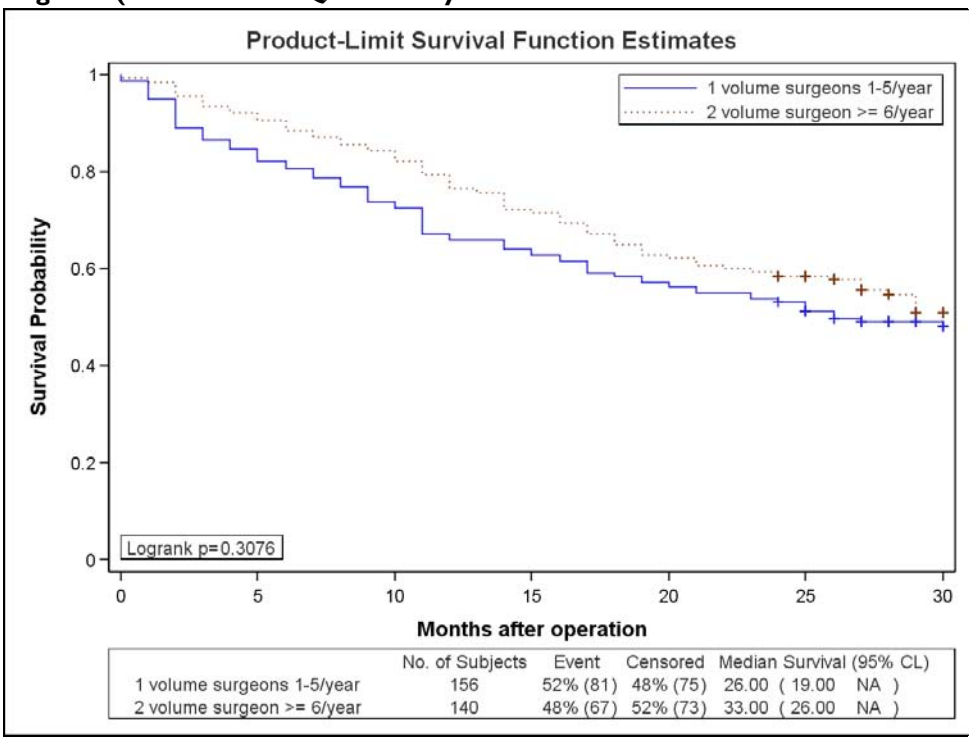


Table 5.13 present results using the AHRQ threshold. As shown in Table 5.13, results of logistic regression (with adjustment for clustering of patients within centres as explained in Chapter 3 on page 30 (section 3.6.1.3) suggest that high volume surgeons have a lower 2-year mortality than low volume surgeons (without reaching statistical significance), after adjustment for patients and tumour characteristics: the odds ratio and 95%CI is 1.30 (0.88, 0.91)

Table 5.13: Oesophageal cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality (surgeon volume)

Model without adjustment for case mix				
Factor		OR	95%CI	
Surgeon volume	1-5/year	1.38	0.92	2.08
	≥ 6 /year	1.00	1.00	1.00
Model with adjustment for patient characteristics (MCD data)				
Surgeon volume	1-5/year	1.24	0.83	1.86
	≥ 6 /year	1.00	1.00	1.00
sex	(male vs female)	1.31	0.61	2.81
age	(increase of 1 y.)	1.02	1.00	1.04
Principal Diagnosis	Oesophagus vs cardia tumour	0.85	0.52	1.40
		1.00	1.00	1.00
Charlson score	Increase of 1	1.28	1.01	1.63
Model with adjustment for patient and tumour characteristics (MCD and BCR data)				
Surgeon volume	1-5/year	1.30	0.88	1.91
	≥ 6 /year	1.00	1.00	1.00
sex	(male vs female)	1.53	0.62	3.75
age	(increase of 1 y.)	1.03	1.01	1.05
Principal Diagnosis	Oesophagus vs cardia tumour	0.70	0.36	1.36
		1.00	1.00	1.00
Charlson score	Increase of 1	1.50	1.16	1.92
Stage	X	1.94	0.74	5.09
	IV	3.66	0.96	13.91
	III	4.80	2.37	9.75
	II	1.05	0.42	2.62
	0-I	1.00	1.00	1.00
Histology	X	0.43	0.14	1.33
	Adenocarcinoma	0.58	0.34	0.99
	Other	1.98	0.61	6.37
	Squamous Cell Carcinoma	1.00	1.00	1.00

Table 5.14 presents the same analysis on the short term mortality (at 3 months). Results are consistent: the data suggest that high volume surgeons have a high mortality than low volume surgeons (using the AHRQ cut off of 6 interventions per year): the odds ratio and 95% CI was 1.31 (0.78, 2.18).

Table 5.14: Oesophageal cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 3-months mortality (surgeon volume)

Model without adjustment for case mix				
Factor		OR	95%CI	
Surgeon volume	1-5/year vs ≥ 6/year	2.54	1.31	4.93
Model with adjustment for patient characteristics (MCD data)				
Surgeon volume	1-5/year vs ≥ 6/year	1.76	0.92	3.35
Model with adjustment for patient and tumour characteristics (MCD and BCR data)				
Surgeon volume	1-5/year vs ≥ 6/year	1.31	0.78	2.18

5.2.7 Discussion

5.2.7.1 External validation of the definition of the procedure

According to the Belgian Cancer Registry, there were 1 140 new cases of oesophagus and cardia cancer in Belgium in 2004 (see Table 5.1 (page 45). In the 2004 Minimal Clinical Data (MCD), we identified 1 401 patients with the principal diagnosis of malignant neoplasm of oesophagus or cardia.

Only 377 of these 1 401 patients (26.9%) underwent major oesophageal cancer surgery. The MCD do not allow a distinction between surgery with palliative and curative intent. This percentage is similar to the one found in other countries.

- In an English study, 30.9% of patients which were identified with a malignant neoplasm of the oesophagus or cardia, underwent surgical resection.[123]
- Among all 4 904 Swedish residents who were diagnosed with oesophageal cancer in the period 1987-2000, only 24.4% underwent oesophageal cancer surgery with curative intent. This study only included resectable oesophageal cancers which implies that oesophagectomies with palliative intent were not counted.[127]

5.2.7.2 Summarized results of literature review

The systematic literature search identified 8 systematic reviews in which the volume outcome association (VOA) for oesophageal cancer surgery was studied.[1, 5, 59, 60, 62, 64, 69, 70] These systematic reviews were based on 22 primary studies.[93, 115, 116, 122, 123, 128-144] An additional search for more recent studies which were published in the period 2004-2009 resulted in 15 additional primary studies.[14, 15, 50, 70, 124, 127, 145-153] This brings us to a total of 37 primary studies that analyzed the volume outcome association for oesophageal cancer surgery. For more insight in how these studies were retrieved see the Supplement.

On the basis of the systematic reviews it was concluded in Chapter 2 (see Table 2.2 on page 19) that there is evidence for an inverse relation between hospital volume and mortality for oesophageal cancer surgery.

This means that mortality decreases when the number of procedures performed by a hospital increases. A similar relationship was found between surgeon volume and mortality.

The highest-quality systematic review on oncologic procedures we identified in the literature study is the one by Killeen and al..[69] The interesting thing about this study is that the authors estimated the number needed to treat, that is the number of patients who must be treated to prevent one adverse outcome, from pooling available absolute risk differences for each procedure. With respect to oesophageal cancer surgery, Killeen and colleagues calculated that the number of oesphagectomies that a high-volume provider needs to prevent one death is as low as seven to nine.

Two minimal hospital volume thresholds were retained from the literature search: lower threshold of 6 oesophagectomies per annum and an upper threshold of 13 oesophagectomies per annum. These thresholds correspond more or less with the thresholds in the systematic reviews. More importantly, they correspond with the thresholds issued by the US Agency for Healthcare Research and Quality (i.e. 6 per annum) and the US Leapfrog Group (i.e. 13 per annum).[29, 126]

Detailed results of the literature review, using the data extraction template, are described in the evidence tables in Appendix 8.

5.2.7.3 *Comparative analysis of literature and Belgian data*

DEFINITION OF VOLUME

In Table 5.10, Belgian hospital volumes for oesophageal cancer surgery were compared with international volume thresholds. It appears that 10 hospitals (14%) attain the threshold issued by the US Agency for Healthcare Research and Quality, which is set at 6 oesophagectomies per annum.[29] The cut off of 13 procedures (in use by the US Leapfrog Group) is only attained by 4 Belgian hospitals (5.7%).[126]

When the same hospital thresholds are applied on Belgian surgeon volumes (because of lack of proper surgeon volume thresholds), we find that 8 surgeons (8.3%) attain the 6/yr threshold and 6 surgeons attain the 13/yr cut off (see Table 5.12).

OUTCOME

In this KCE study, in-hospital and approximate 30-day mortality is 9.1% after oesophagectomy for cancer. One-year and two-year mortality rates are, respectively, 30% and 45%. Median survival time was 29 months. Table 5.15 compares these Belgian outcome data with those published in the studies that were selected.

- In the US study by Birkmeyer et al. for the years 1994 through 1999, the observed mortality rate (in-hospital or within 30 days after surgery) varies from 23.1% in very low volume hospital (<2/yr) to 8.1% in very high volume hospital (>19/yr). Unfortunately, Birkmeyer does not mention the mortality over all hospitals.[115]
- A retrospective review of 1 125 English patients who had surgery for cardio-oesophageal cancer between 1992 and 1996 showed an overall 30-day mortality rate of 10.0% with a median survival of 14 months and a 5-year survival rate of 17.2%. Survival rate after one, two and three years can be deduced from Figure 1 of the article and is respectively 54.8, 31.7 and 23.9%.[123]
- Canadian administrative data from between 1994 and 1999 showed a 30-day mortality of 13.4% after oesophagectomy.[142]
- The Swedish study by Rouvelas et al. which included all Swedish residents with resectable oesophageal cancer in the period 1987-2000 found a 30-day mortality of 7.0% after oesophagectomy with curative intent. The observed 1-year, 3-year and 5-year survival rate were respectively 56.1%, 32.0% and 25.2%.[127] Another study by Rouvelas and colleagues studied oesophageal cancer surgery in a more recent period (2001-2005). This study illustrates the improved survival in recent years: 30-day mortality was only 3.0% while 90-day mortality was 7.9%.[148] A very similar patient population was studied in the study carried out by Rutegard et al. where postoperative mortality within 30 days of surgery amounted to only 2.8%.[124]
- Clinical data from 903 oesophageal resections performed in 12 Dutch hospitals between 1990 and 1999 showed an in-hospital mortality of 8.0%. Median survival time was 21 months in low-volume hospitals and 22 months in high-volume hospitals.[152]

This limited comparison leads us to conclude that the 9.1% approximate 30-day mortality found in this study is very comparable to the one found in other studies. Except for the Swedish mortality that is much lower. In relation to one- and two-year survival, this study's results seem higher than those seen elsewhere. A possible explanation for this difference could be that patient selection (i.e. eligibility for surgery) is more selective in Sweden.

Table 5.15: Oesophageal cancer surgery: Mortality: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Birkmeyer 2002 [115]	Gillison 2002 [123]	Urbach 2004 [142]	Rouvelas 2007 [127]	Rouvelas 2007 [148]	Rutegard 2009 [124]	Wouters 2008 [152]
Country Study period	Belgium 2004	USA 1994-1999	England 1992-1996	Canada 1994-1999	Sweden 1987-2000	Sweden 2001-2005	Sweden 2001-2005	NL 1990-1999
In-hospital or 30-day mortality	9.1%	8.1% to 23.1%	10.0%	13.4%	7.0%	3.0%	2.8%	8.0%
Median survival time	29 months		14 months					21 to 22 months
1-year survival	70.2%		54.8%		56.1%			
2-year survival	55.3%		31.7%					

PATIENT CASE MIX

As shown in Table 5.16, the age and sex distributions are similar to those in foreign studies.

Table 5.16: Oesophageal cancer surgery: Patient characteristics: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Dimick 2001 [122]	Gillison 2002 [123]	Rouvelas 2007 [148]	Wouters 2008 [152]
Country **Study period**	Belgium 2004	USA 1984-1999	England 1992-1996	Sweden 2001-2005	Netherlands 1990-1999
Age	64 years (mean)	61 years (mean)	66 years (median)	66 years (mean)	65 years (mean)
Male	80.1%	74%	70.9%	80.6%	76.1%

A comparison of tumour characteristics in Table 5.17 shows a higher rate of oesophagus cancers among the patients in the KCE study; almost 70% versus approximately 55% in other European studies. Tumour histology, on the other hand, seems very similar to those in other studies. The tumour stage distribution, finally, is very similar to the one in the Swedish registry.[148] English patients had a more advanced stage of the cancer while the Dutch study had more patients with stage 2 disease.[123, 152]

Table 5.17: Oesophageal cancer surgery: Tumour characteristics: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Gillison 2002[123]	Rouvelas 2007[148]	Wouters 2008[152]
Country	Belgium	England	Sweden	Netherlands
Study period	2004	1992-1996	2001-2005	1990-1999
Tumour location				
Oesophagus	69.3%	58.1%	54.0%	52.9%
Cardia	30.7%	41.9%	46.0%	46.2%
Tumour histology				
Squamous cell carcinoma	27.4%	26.3%	24.6%	32.0%
Adenocarci-noma	67.1%	70.6%	74.0%	64.8%
Other	5.5%	3.1%	1.4%	3.2%
Tumour stage				
I	22.1%	6.3%	18.9%	11.5%
II	32.5%	31.7%	29.5%	41.6%
III	37.2%	50.1%	40.4%	32.5%
IV	8.2%	11.8%	11.2%	12.7%

In the KCE study, patients having surgery at high- and low-volume hospitals were very similar with respect to age, gender and Charlson score (see Table 5.9). Most other studies observed the same patient similarity across the volume strata.[122, 123, 149] Rouvelas et al., however, observed a slightly lower proportion of patients without any co morbidity in the low-volume group.[148] Birkmeyer et al., on the other hand, saw that patients at lower volume hospitals tended to have more co morbidities.[153]

With respect to the distribution of the tumour stage among the hospital categories, it was very difficult to find a certain pattern in the BCR data (see Table 5.9). Other studies mentioned that they did not see a material difference in tumour stage between the volume strata.[148, 149, 153] Gillison et al., on the other hand, saw that low-volume surgeons had more advanced stage patients.[123] Finally, Wouters et al. observed that more patients from the high-volume hospitals had an advanced stage of the disease.[152] Note that there was only one US study in which tumour characteristics were applied for risk adjustment i.e. the one by Birkmeyer published in 2007.[153]

The following characteristics were additionally used for risk adjustment in other studies but could not be applied in the KCE study because they are not available in MCD or BCR data:

- race;[115, 153]
- median household income;[115, 153]
- college education;[153]
- acuity of index admission i.e. elective, urgent or emergent;[115, 123, 153]
- surgical approach i.e. transthoracic or transhiatal;[124, 152]
- surgical anastomosis i.e. cervical, thoracic, abdominal;[148, 152]
- proximal section margin from the tumour i.e. <10, 10-50, >50 mm;[148]
- use of (neo)adjuvant therapy i.e. chemotherapy or radiation therapy;[123, 124, 127, 148, 152, 153]
- intention of surgery i.e. palliative or curative.[123, 148]

As mentioned, several authors adjusted for use of neoadjuvant (given before the primary treatment) or adjuvant (given after the primary treatment). Most of them found that the use of adjuvant radiation therapy and chemotherapy did not vary systematically by hospital volume.[124, 127, 148, 153] In addition, Birkmeyer et al. assessed that adjusting for differences in the use of this therapy only had a negligible effect in attenuating differences in 5-year survival.[153]

Wouters et al., on the other hand, found that (neo) adjuvant treatment varied widely between the Dutch volume groups. However, these factors were not significantly related to mortality.[152]

VOLUME OUTCOME ASSOCIATION

The present study did not find an inverse relationship between hospital volume and 2-year mortality. This result contrasts with the findings from the literature review which concluded that there is evidence for an inverse relationship between hospital volume and mortality for oesophageal cancer surgery.

The question rises whether this discrepancy between our findings and literature has something to do with the missing data on stage. The fact that mortality rate was not substantially higher in the patients whose stage was missing in the BCR (see Figure 5.3) seems to indicate, however, that these patients are randomly divided into the four disease stages. Ideally, though, this assumption should be checked with the help of sensitivity analyses, which was not done due to time constraints. In addition, we noticed that many hospitals – low-volume as well as high-volume – missed data on stage and that the percentage of missing data varied among these hospitals (see Figure 5.1). Despite the failure to retrieve information on disease stage, this problem did not restrain us from drawing conclusions on the volume outcome association.

In 2003, Birkmeyer et al. proved that 46% of the apparent effect of hospital volume was actually attributable to surgeon volume.[116] When we consider this possibility, a possible explanation for the fact that low-volume hospitals have such low mortality could be that these patients were treated by high-volume surgeons. This hypothesis was tested, however, and rejected since we found that patients at low-volume hospitals were operated on by low-volume surgeons.

With respect to surgeon volume, the data suggest that there is an inverse association with 3 month mortality, although it did not reach statistical significance. 3 months mortality was 13.5% for surgeons with less than 6 interventions per year and 6.4% for surgeons with at least 6 interventions per year. Results at two years were consistent. Several years of observations are required to increase precision of effects.

Key points on volume outcome association for <u>oesophageal</u> cancer surgery

- **A total of 1 401 patients were hospitalized in 2004 (retrieved in MCD 2004) with a diagnosis of malignant neoplasm of oesophagus or cardia. 27% of those patients underwent a resection of the oesophagus.**
- **The population studied consisted of those 329 patients with oesophagectomy and whose data could be linked to IMA databases.**
- **In 80% of the cases, information on tumour could be retrieved in the BCR database. Data on stage was available for 68% of stays. Low-volume as well as high-volume hospitals missed data on stage; there seemed to be no association between percentage of linkage with BCR and hospital volume.**
- **These interventions were performed in 72 centres by 99 surgeons. 10 hospitals and 8 surgeons had a volume higher or equal to 6 interventions per year (the current AHRQ criteria in US).**
- **Two-year mortality was 45%. Regression models were fitted to assess the association between hospital or surgeon volume with this outcome. The following factors were taken into account in all analyses: sex, age, principal diagnosis (oesophagus or cardia), Charlson score (co morbidity), tumour stage and tumour histology (adenocarcinoma or squamous cell carcinoma).**
- **Based on the study of systematic reviews, it was concluded that there is evidence for an inverse relation between hospital volume and mortality for oesophageal cancer surgery. Two minimal hospital volume thresholds were retained from the literature search: a lower threshold of 6 oesophagectomies per annum and an upper threshold of 13 oesophagectomies per annum.**
- **In contrast to scientific literature, however, Belgian data did not show an inverse relationship between volume of centres and 2-year mortality: respectively 43.8% in centre with less than 6 interventions per year and 45.4% in centres with at least 6 interventions per year.**
- **The literature review also concluded that there is inverse relationship between surgeon volume and mortality.**
- **As well, Belgian data suggest an inverse association (not statistically significant) between the volume of surgeons and 3 months mortality: 13.5% for surgeons with less than 6 interventions per year and 6.4% for surgeons with at least 6 interventions per year. Results at two years were consistent.**
- **Several years of observations are required to increase precision of effects.**
- **Because data were not retrieved or not available in the databases, the following characteristics could not be used for risk adjustment: use of (neo) adjuvant therapy (chemo- or radiation therapy), acuity of admission (elective versus urgent), intention of surgery (palliative versus curative).**

5.3 PANCREATIC CANCER SURGERY

5.3.1 General description of procedure

Surgical removal of the pancreas (i.e. pancreatectomy) is still the only chance of long-term survival for patients with pancreas cancer. Belgian guidelines, which were recently published in KCE report 105, recommend that "patients with resectable pancreatic cancer who are fit for surgery should undergo radical pancreatic resection (pancreaticoduodenectomy for pancreatic head tumours and distal pancreatectomy for pancreatic body and tail tumours)".[154]

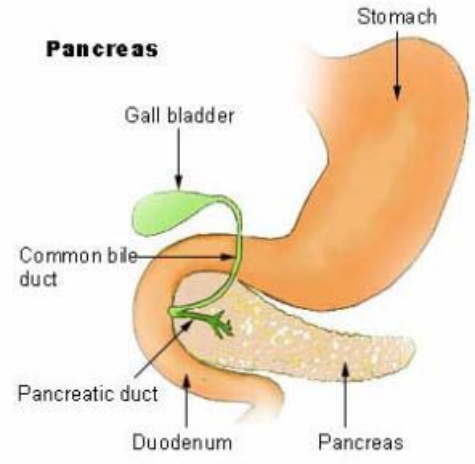
5.3.2 Primary data selection in MCD 2004

First selection of data was based on ICD-9-CM diagnostic and procedure codes as proposed by Glasgow et al., and on NIHDI procedure codes.[155]

In addition to the ICD-9-CM diagnosis code for pancreatic cancer (157), codes for malignant neoplasm of duodenum (1520), of extrahepatic bile ducts (1561) and of ampulla of Vater (1562) were added to the selection criteria.

The rationale of this elaborate selection is that patients can be labelled with the latter diagnoses at the end of their hospital stay - and be coded as such in the MCD - on the basis of suspicious lesions that were visualised with medical imaging. It is often only a few days after the hospitalisation that the final diagnosis can be attributed on the basis of the results of the histopathologic examination of the tumour. Therefore, it was decided to make a broad selection at the level of MCD.

Figure 5.8: Pancreas and surrounding organs



Source: Wikimedia Commons (Public Domain)

This selection resulted in a total of 4 949 stays (Table 5.18). This selection is too broad since it also includes patients with a secondary diagnosis of malignant neoplasm. This implies that patients with pancreatic cancer as pre-existing condition are also included (see section 4.2.1, page 39). The definition of the population studied (i.e. patients with a principal diagnosis of malignant neoplasm of pancreas who underwent surgical resection of the pancreas) will be defined in the next section.

Table 5.18: Pancreatic cancer surgery: Primary data selection in MCD 2004

SELECTION 1	SELECTION 2	SELECTION 3
ICD-9-CM (principal and secondary) diagnosis code	**ICD-9-CM procedure code**	**NIHDI procedure code**
1520 Malignant neoplasm of duodenum	5251 Proximal pancreatectomy	242023 Duodeno-pancreatectomy.
1561 Malignant neoplasm of extrahepatic bile ducts	5252 Distal pancreatectomy	242045 Left-hemipancreatectomy with pancreatico-jejunal anastomosis, or as good as total pancreatectomy ((95%).
1562 Malignant neoplasm of ampulla of Vater	5253 Radical subtotal pancreatectomy	242060 Left-hemipancreatectomy or enucleation of pancreatic tumour or extirpation of a pancreatic sequestrum.
157x Malignant neoplasm of pancreas	5259 Other partial pancreatectomy	
	526 Total pancreatectomy	
	527 Radical pancreatico-duodenectomy	
STAYS SELECTED = 4 640 stays	**STAYS SELECTED = 593 stays**	**STAYS SELECTED = 560 stays**
TOTAL STAYS SELECTED= 4 949 stays (selection 1 OR selection 2 OR selection 3)		

5.3.3 Definition of procedure

5.3.3.1 Primary hospital stays in Minimal Clinical Data

2 531 of the 4 949 selected hospital stays had a principal diagnosis of malignant neoplasm of duodenum (4.3% i.e. 109 stays), of extrahepatic bile ducts (7.4% i.e. 186 stays), of ampulla of Vater (4.4% or 110 stays) or of pancreas (84.0% i.e. 2 126 stays). As shown in Table 5.19, only 322 stays (12.7% of 2 531 stays) corresponding to 320 patients (17.4% of 1 824 patients) had a procedure for peripancreatic cancer surgery as described above.

Table 5.19: Pancreatic cancer surgery: Principal diagnosis and percentage of cancer surgery (per stay and per patient)

Principal diagnosis	Total number hospital stays	Stays with pancreatic cancer surgery		Total number patients	Patients with pancreatic cancer surgery	
		Nbr	%		Nbr	%
1520 Malignant neoplasm of duodenum	109	24	22.0	87	24	27.6
1561 Malignant neoplasm of extrahepatic bile ducts	186	16	8.6	146	16	11.0
1562 Malignant neoplasm of ampulla of Vater	110	36	32.7	85	36	42.4
157 Malignant neoplasm of pancreas	2 126	246	11.6	1 526	244	15.99
Total	2 531	322	12.7	1 842	320	17.37

The APR-DRGs of these 322 stays are described in Table 5.20.

To create a study population of patients which is as homogenous as possible, we retained only the 311 stays (corresponding to 309 patients) that are included in the APR-DRGs 220 "Major stomach, oesophageal & duodenal procedures" and 260 "Pancreas, liver & shunt procedures".

Table 5.20: Pancreatic cancer surgery: All APR-DRGs of stays with a principal diagnosis of a malignant neoplasm of duodenum, of extrahepatic bile ducts or of pancreas AND with surgical resection of the pancreas

APR-DRG	Number	Percent
004-TRACHEOSTOMY EXCEPT FOR FACE, MOUTH & NECK DIAGNOSES / p3 - P	2	0.62
220-MAJOR STOMACH, ESOPHAGEAL & DUODENAL PROCEDURES / 6 - P	**19**	**5.90**
221-MAJOR SMALL & LARGE BOWEL PROCEDURES / 6 - P	1	0.31
223-MINOR SMALL & LARGE BOWEL PROCEDURES / 6 - P	1	0.31
229-OTHER DIGESTIVE SYSTEM PROCEDURES / 6 - P	3	0.93
260-PANCREAS, LIVER & SHUNT PROCEDURES / 7 - P	**292**	**90.68**
261-MAJOR BILIARY TRACT PROCEDURES / 7 - P	3	0.93
264-OTHER HEPATOBILIARY & PANCREAS PROCEDURES / 7 - P	1	0.31
TOTAL	322	100.00

Results of definition: 311 stays with pancreatic cancer surgery were selected in the Minimal Clinical Data. These stays concerned 309 individual patients and 74 centres.

5.3.3.2 Linkage with data from IMA and BCR

First, the Minimal Clinical Data (MCD) were linked with the Common Sickness Funds Agency (IMA) database. As shown in Table 5.21, 96.8% of 311 MCD stays could be linked with IMA data.

Second, 79.8% of the original 311 MCD stays were linked to BCR data on the basis of ICD-10 codes C17.0, C24.0, C24.1 and C25. Originally, we planned to limit the data to those records with a histopathologically confirmed pancreatic cancer (ICD-10 code C25).

However, as shown in Table 5.22, this would have resulted in a substantial decrease in number of records since only 188 MCD stays with ICD-9-CM code 157 could be linked with BCR data with ICD-10 code C25. A second reason to include all peripancreatic cancer diagnoses instead of limiting the data to the pancreas cancers only is the consistency between the diagnoses in the two databases.

Table 5.22 illustrates, for example, how 13 out of 16 stays in MCD with diagnosis "malignant neoplasm of extrahepatic bile ducts" correspond with the same diagnosis in the BCR. This proofs that our original concern that the final diagnosis (based on the histopathologic examination of the tumour) would be different from the clinical diagnosis encoded in the MCD, was unfounded. Because of our broad selection in relation to the type of cancer that is surgically treated with a pancreatectomy, it is more correct to label these cancers as peripancreatic cancers.

Third, information on tumour stage in the BCR was available for 207 patients with peripancreatic cancer. These account for 66.6% of the initial 311 stays in the MCD selection. This percentage varies strongly among the hospitals as is shown in Figure 5.9 where the red columns indicate the number of procedures for which stage is missing.

Table 5.21: Pancreatic cancer surgery: Percentage of linkage of MCD data with data from IMA and BCR

	Number	%
Number of stays in MCD selection	311	100
Linkage with IMA	301	96.8
Linkage with BCR	248	79.8
Linkage with BCR and data about stage	207	66.6

Table 5.22: Pancreatic cancer surgery: Linkage of principal diagnosis in MCD and diagnosis in BCR

	Diagnosis in BCR (ICD-10)					
Principal diagnosis in MCD (ICD-9-CM)	C17.0 Malignant neoplasm of duodenum	C24.0 Malignant neoplasm of extrahepatic bile ducts	C24.1 Malignant neoplasm of ampulla of Vater	C25 Malignant neoplasm of pancreas	Missing in BCR	Total in MCD
1520 Malignant neoplasm of duodenum	8 42.1%	1 5.3%	5 26.3%	2 10.5%	3 15.8%	19 100%
1561 Malignant neoplasm of extrahepatic bile ducts	0 0.0%	13 81.3%	0 0.0%	0 0.0%	3 18.8%	16 100%
1562 Malignant neoplasm of ampulla of Vater	1 2.9%	0 0.0%	22 62.9%	3 8.6%	9 25.7%	35 100%
157 Malignant neoplasm of pancreas	0 0.0%	2 0.8%	8 3.3%	183 75.9%	48 19.9%	241 100%
Total in MCD	9 2.9%	16 5.1%	35 11.3%	188 60.5%	63 20.3%	311 100%

Results of linkage: 248 patients with peripancreatic cancer surgery could be linked with Belgian Cancer Registry data. For 207 of these patients, the BCR contained information on the stage of the tumour.

5.3.3.3 Patient and tumour characteristics

Peripancreatic cancer contains following tumour locations (ICD-10 codes):

- C17.0 for malignant neoplasm of duodenum
- C24.0 for malignant neoplasm of extrahepatic bile ducts
- C24.1 for malignant neoplasm of ampulla of Vater
- C25 for malignant neoplasm of pancreas.

Tumour histology is not used as a variable for risk adjustment and is therefore not analyzed in this section. For a detailed description of TNM classification and TNM stages for pancreatic cancer, we refer to KCE report 105.[154]

Table 5.23 shows that the mean age of the 311 patients selected in the MCD was 65 years (median age is 66 years) and 55% of them were men. Three quarters (75.8%) of peripancreatic tumours were located in the pancreas. Majority of patients had stage 2 pancreatic cancer (56%).

Table 5.23: Pancreatic cancer surgery: Patient and tumour characteristics

		Number	%
Age (years)			
	mean/median	65.0/66	
	Std	10.	
	Total	*311*	
Sex			
	Male	171	54.98
	Female	140	45.02
	Total	*311*	
Tumour location			
	Duodenum	9	3.63
	Extrahepatic bile ducts	16	6.45
	Ampulla of Vater	35	14.11
	Pancreas	188	75.81
	Total	*248*	
Tumour stage			
	I	46	22.22
	II	116	56.04
	III	20	9.66
	IV	25	12.08
	Total	*207*	

5.3.4 Definition of volume

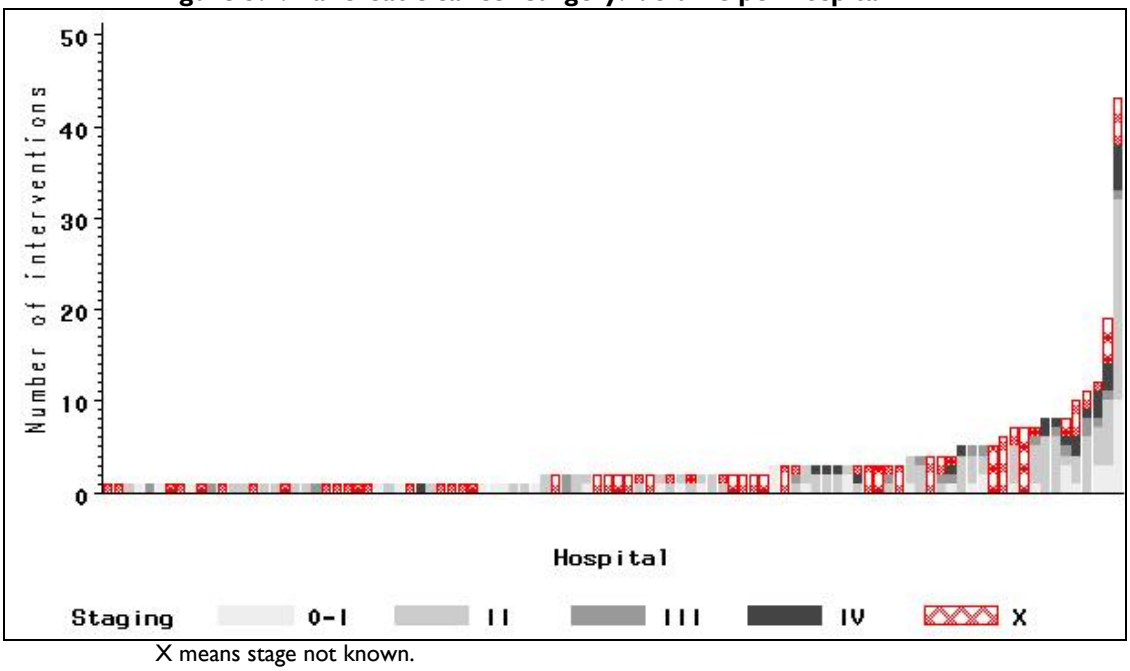
As shown in Table 5.24, the 311 procedures (selected in the MCD) were performed in 74 centres by 112 surgeons. Three quarters of the hospitals performed 4 procedures or less.

Table 5.24: Pancreatic cancer surgery: Summary measures of volume per hospital and per surgeon

	Number	mean	minimum	25th Pctl	50th Pctl	75th Pctl	maximum
Hospitals	74	4.2	1.0	1.0	2.0	4.0	50.0
Surgeons*	112	2.7	1.0	1.0	1.0	2.0	25.0

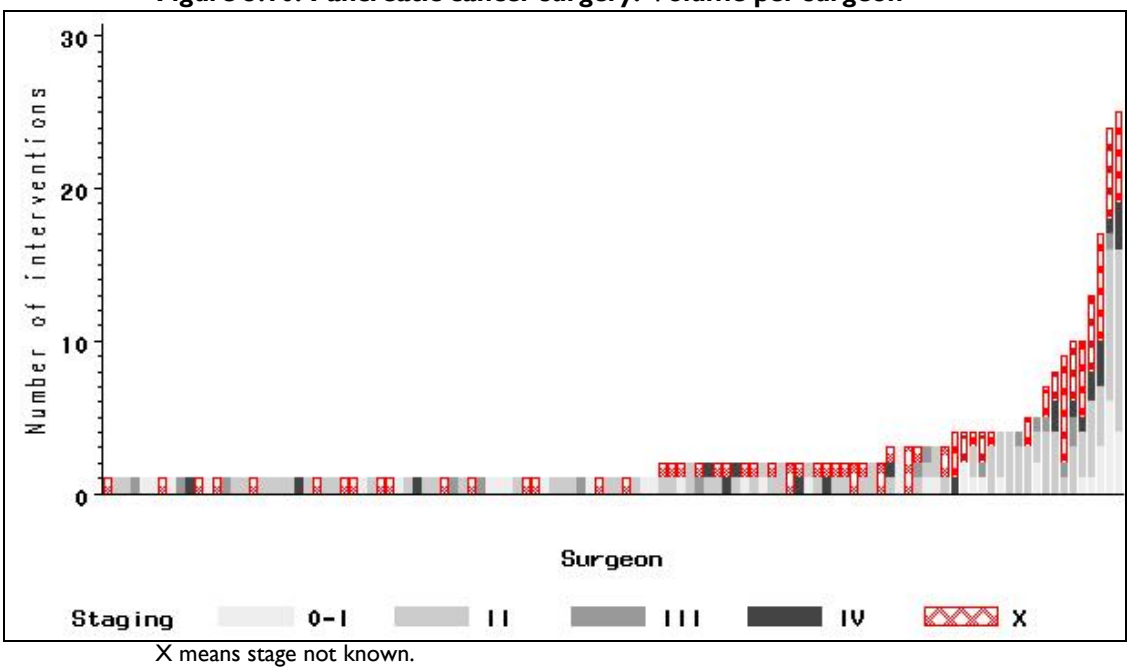
* For 14 procedures information on surgeon was missing.

Figure 5.9: Pancreatic cancer surgery: Volume per hospital



X means stage not known.

Figure 5.10: Pancreatic cancer surgery: Volume per surgeon



X means stage not known.

5.3.5 Definition of outcomes

5.3.5.1 Mortality

Table 5.25 shows that in-hospital and approximate 30-day mortality is 8.3% after pancreatic cancer surgery (25 deaths). Survival 1 year after surgery is 68.1%. After 2 years, survival is only 46.8%. Median survival time was 23 months (not visible in Table 5.25).

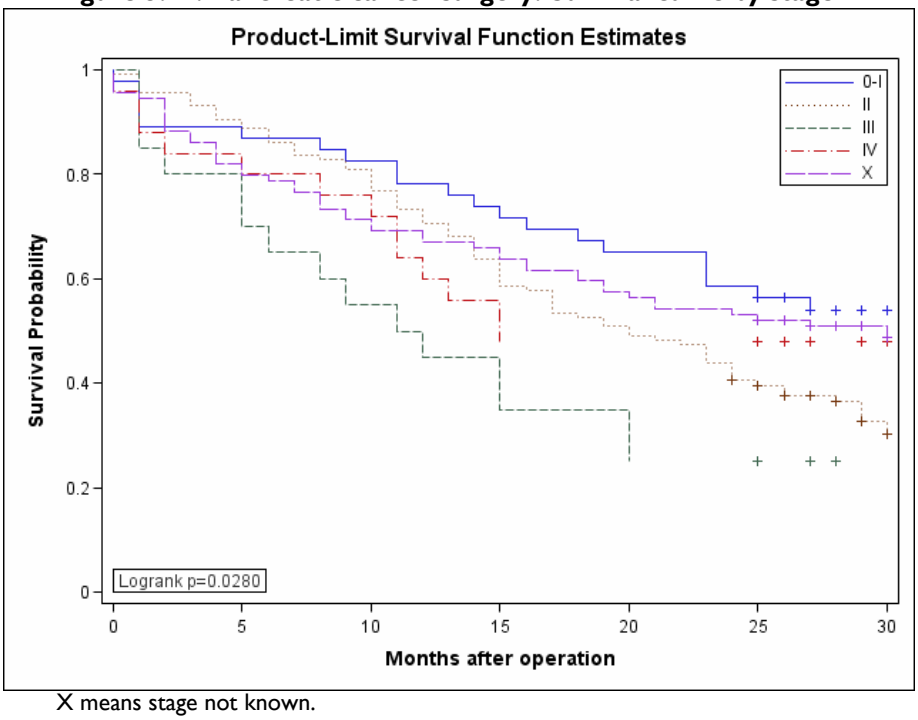
Table 5.25 illustrates well how 2-year survival after pancreatectomy is affected by the principle diagnosis. Patients who are diagnosed with a malignant neoplasm of extrahepatic bile ducts or ampulla of Vater have a far better survival rate (66.7%) than those with a malignant neoplasm of duodenum (36.8%) or pancreas (43.3%). These differences confirm the necessity to add the principal diagnosis to the risk adjustment.

Table 5.25: Pancreatic cancer surgery: Mortality results

	Number	Number deaths	% deaths	Survival
In-hospital and approximate 30-day mortality	301	25	8.3	
3 months mortality	301	34	11.3	
1-year mortality	301	96	31.9	68.1
2-year mortality	301	160	53.2	46.8
2-year mortality in function of principal diagnosis:				
152 Malignant neoplasm of small intestine, including duodenum	19	12	63.2	36.8
1561 Malignant neoplasm of extrahepatic bile ducts or 1562 Malignant neoplasm of ampulla of Vater	51	17	33.3	66.7
157 Malignant neoplasm of pancreas	231	131	56.7	43.3

Figure 5.11 illustrates how survival improves with decreasing disease stage. Patients with unknown disease stage (purple line) are in between those with stage 1 and stage 2 disease and have a rather good survival rate.

Figure 5.11: Pancreatic cancer surgery: Survival curve by stage



X means stage not known.

Because of lack of reliable short term mortality data, analyses on volume outcome relationship are limited to 2-year mortality. This implies that additional outcome measures such as postoperative bleeding rate and infection rate, which are listed in Table 2.1 (page 16), are not analyzed in this report.

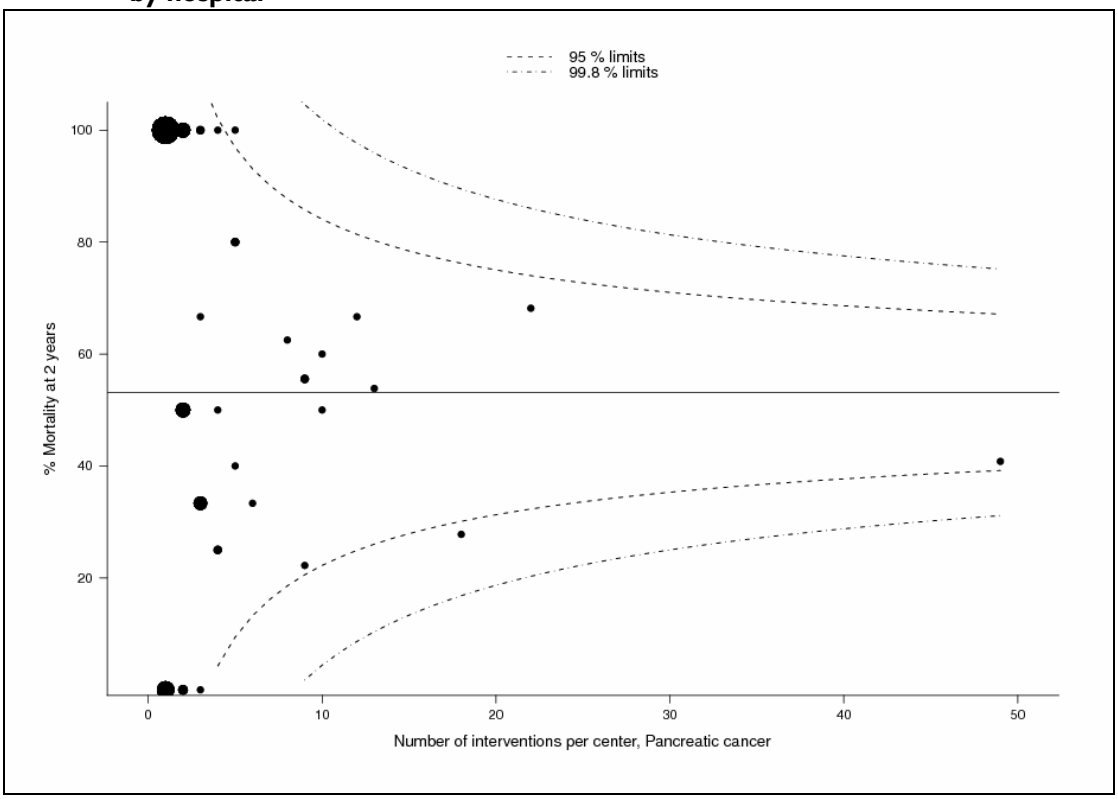
5.3.6 Volume outcome relationship

5.3.6.1 Analysis by hospital

MORTALITY RATE, BY HOSPITAL

Figure 5.12 presents the funnel plot of the number of pancreatectomies per centre and the 2-year mortality (observed, i.e. without risk adjustment). The size of the point is proportional to the number of hospitals with the same volume and the same outcome. The horizontal line represents the overall 2-year mortality i.e. 53.2%. One low volume centre is just above the 95% limits of variability, and one medium volume centre has better performance. None of the hospitals are outliers above the 99.8% limits of variability, but one medium volume hospital shows better outcome (low limits of 95% variability).

Figure 5.12: Pancreatic cancer surgery: Funnel plot of 2-year mortality rate, by hospital



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

Because the majority of centres were below the international cut off of 11 interventions/year, 3 volume categories were defined below this cut off and two above. Table 5.26 shows that 2-year mortality is 65% in very low-volume hospitals (1-2 procedures/yr) and less than 50%. in hospitals with at least 6 procedures per annum. Patients who are treated in hospitals with less than 6 interventions per year are 67 years old while patients treated in bigger hospitals are from 63 to 34 years old. The percentage of patients with Charlson score ≥ 3 is 21% in very low-volume hospitals, compared to 5.6% in the largest hospitals (>20/year). The highest volume hospitals (>20/yr) have operated 25% of patients with malignant neoplasm of extrahepatic bile ducts or ampulla of Vater (which have a better survival rate as seen in Table 5.25).

Table 5.26: Pancreatic cancer surgery: Differences in case-mix and outcomes by hospital volume

		Hospital volume					
		1-2/yr	3-5/yr	6-10/yr	11-20/yr	>20/yr	All
Number centres		41	20	7	3	2	73
Number stays		57	69	61	43	71	301
Gender	%male	56.1	52.2	55.7	58.1	57.7	55.8
	Mean	66.8	66.9	62.9	63.3	64.2	64.9
Age	Std	9.9	9.7	11.5	10.0	10.6	10.5
Charlson ≥ 3	%	21.1	11.6	16.4	4.7	5.6	12.0
	Mean	1.1	1.0	1.2	0.5	0.7	0.9
Charlson	Std	1.6	1.5	1.7	0.9	1.1	1.4
Severity index (APR-DRG)							
1	%	1.8	1.4	3.3	0.0	2.8	2.0
2	%	7.0	5.8	3.3	14.0	4.2	6.3
3	%	26.3	14.5	26.2	16.3	32.4	23.6
4	%	64.9	78.3	67.2	69.8	60.6	68.1
Principal Diagnosis (MCD)							
Duodenum tumour	%	5.3	8.7	3.3	7.0	7.0	6.3
Extrahepatic bile ducts or ampulla of Vater tumour	%	14.0	14.5	16.4	11.6	25.4	16.9
Pancreas tumour	%	80.7	76.8	80.3	81.4	67.6	76.7
Length of stay	Median	25.0	24.0	23.0	27.0	18.0	22.0
In-hospital + approximate 30-day mortality	%	10.5	8.7	11.5	0.0	8.5	8.3
2-year mortality	%	**64.9**	**55.1**	**49.2**	**46.5**	**49.3**	**53.2**
Missing data in BCR (stage)	%**	35.1	21.7	37.7	39.5	26.8	31.2
Stage							
I	%	32.4	20.4	15.8	15.4	25.0	22.2
II	%	51.4	61.1	57.9	50.0	55.8	56.0
III	%	13.5	9.3	13.2	11.5	3.8	9.7
IV	%	2.7	9.3	13.2	23.1	15.4	12.1

** Missing data for stage are not included in % for distribution of stays across stage

RELATION BETWEEN HOSPITAL VOLUME AND 2-YEAR MORTALITY RATE

Table 5.27 shows the results analyses which were performed with the international cut off of 11 procedures per year (AHRQ and Leapfrog threshold).[29, 126]

Table 5.27: Pancreatic cancer surgery: mortality per hospital based on international volume thresholds

International cut off	Hospital volume category	Number centres	Number patients	3-months mortality		2-year mortality	
				Number deaths	%	Number deaths	%
AHRQ cut off 11/year	1-10/year	68	187	23	12.3	105	56.1
	≥11/year	5	114	11	9.6	55	48.2

Table 5.28 shows results from the logistic regression model (with GEE to adjust for clustering of patients within hospitals) used to examine the relation between hospital volume and 2-year mortality following peripancreatic cancer surgery.

The first model was performed without adjustment for case mix and suggest that the odd of 2 year mortality is 30% higher in hospitals with less than 11 interventions per year than in hospital with at least 11 interventions per year: odds ratio and 95% CI: 1.31 (0.71, 2.42) This result did not reach statistical significance.

In the second model, an adjustment was made for patient characteristics which were retrieved from the MCD data i.e. sex, age, principal diagnosis and Charlson score. This adjustment does not affect the results of the volume effect (odss ratio and 95% CI 1.29 (0.85, 1.94)) Patients with a duodenum tumour have a higher mortality than those with a pancreas tumour, although not statistically significant (OR = 1.23, 95%CI 0.40-3.79). Patients with a malignant neoplasm of extrahepatic bile ducts or ampulla of Vater have a significant better survival rate than those with a malignant neoplasm of pancreas (OR = 0.37, 95%CI 0.20-0.69).

The third model also adjusted for tumour characteristics such as stage and histology of the peripancreatic cancer. This adjustment barely affects the effect of volume: OD and 95% CI 1.25 (0.83, 1.89).

Table 5.28: Pancreatic cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality

Model without adjustment for case mix				
Factor		OR	95%CI	
Hospital Volume	1-10/year vs ≥11/year	1.31	0.71	2.42
Model with adjustment for patient characteristics (MCD data)				
Hospital Volume	1-10/year vs ≥11/year	1.29	0.85	1.94
Sex	male vs female	1.24	0.79	1.95
Age	increase of 1 yr	1.03	1.00	1.06
Principal Diagnosis	Duodenum vs Pancreas	1.23	0.40	3.79
	Extrahepatic bile ducts or Ampulla of Vater versus Pancreas	0.37	0.20	0.69
Charlson score		1.07	0.80	1.43
Model with adjustment for patient and tumour characteristics (MCD and BCR data)				
Hospital Volume	1-10/year vs ≥11/year	1.25	0.83	1.89
Sex	male vs female	1.37	0.85	2.20
Age	increase of 1 yr	1.04	1.00	1.08
Principal Diagnosis	Duodenum vs Pancreas	1.06	0.36	3.16
	Extrahepatic bile ducts or Ampulla of Vater versus Pancreas	0.35	0.16	0.77
Charlson score	increase of 1	1.08	0.80	1.45
Stage	unknown	1.31	0.57	3.05
	IV	1.85	0.57	5.97
	III	6.09	1.83	20.20
	II	2.22	1.03	4.76

Sensitivity analyses with the log of volume as predictor showed effects borderline statistically significant (results not shown).

5.3.6.2 *Analysis by surgeon*

MORTALITY RATE, BY SURGEON

Figure 5.13: Pancreatic cancer surgery: Funnel plot of in-hospital and approximate 90-day mortality rate, by surgeon

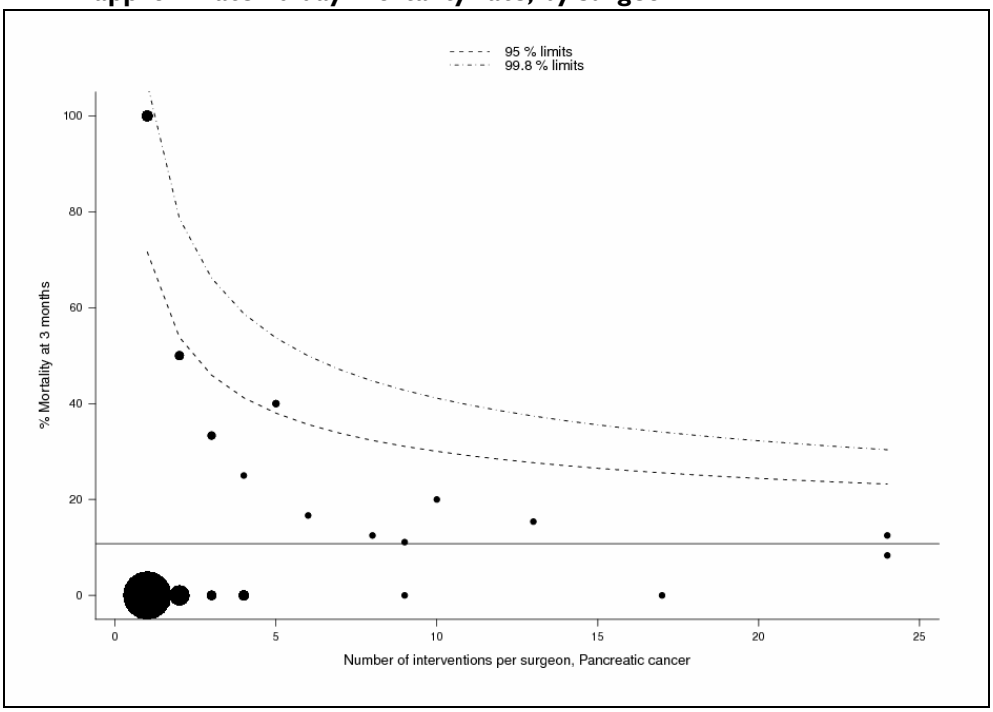
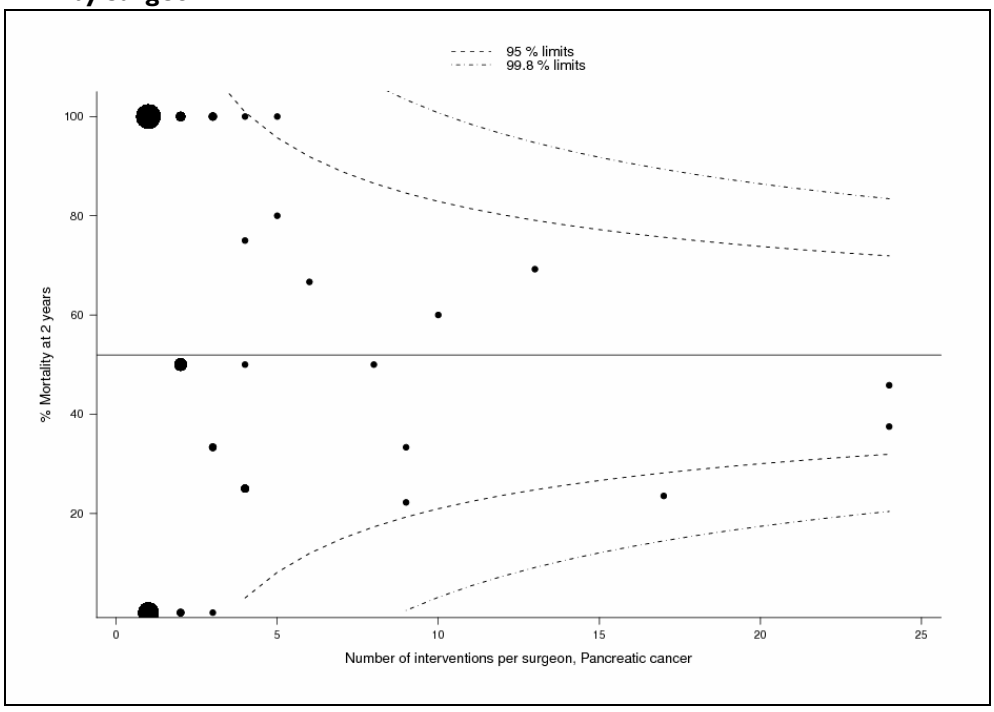


Figure 5.14 presents the funnel plot of the number of pancreatectomies per surgeon and the 2-year mortality (observed, i.e. without risk adjustment). The size of the point is proportional to the number of surgeons with the same volume and the same outcome. One surgeon is an outlier below the 95% limits of variability (with better outcome), and another surgeon with low volume is outlier above the 95% limits of variability.

Figure 5.14: Pancreatic cancer surgery: Funnel plot of 2-year mortality rate, by surgeon



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE SURGEON

The majority of surgeons are below the international cut off of 6 interventions per year. Three categories were defined below that cut off (1/year: 60 surgeons, 2/year and 3-5/year). Table 5.29 show that surgeon performing at least 6 pancreatic cancer have a 2-year mortality rate lower than 50%, while surgeons who perform less than 6 procedures per year have a 2-year mortality rate of above 55%. Patients who are treated by low-volume surgeons seem a bit older and have more co morbidities (i.e. more patients with Charlson score ≥ 3). The highest volume surgeons (>10 procedures/year) have operated more patients with malignant neoplasm of extrahepatic bile ducts or ampulla of Vater (which have a better survival rate as seen in Table 5.25). Finally, distribution of the tumour stages among the surgeon categories does not seem to be related to the volume.

Table 5.29: Pancreatic cancer surgery: Differences in case-mix and outcomes by surgeon volume

		Surgeon volume					
		1/yr	2/yr	3-5/yr	6-10/yr	>10/yr	All
Number surgeons		60	23	17	5	4	109
Number stays		60	45	62	42	78	287
Gender	%male	51.7	55.6	56.5	61.9	56.4	56.1
	Mean	66.9	65.0	66.5	63.2	63.3	65.0
Age	Std	8.6	11.3	9.5	12.2	10.7	10.4
Charlson ≥ 3	%	18.3	17.8	12.9	11.9	2.6	11.8
Charlson	Mean	1.2	1.0	1.1	0.9	0.6	0.9
	Std	1.6	1.5	1.4	1.6	0.9	1.4
Severity index (APR-DRG)							
1	%	1.7	2.2	3.2	2.4	1.3	2.1
2	%	10.0	6.7	3.2	9.5	2.6	5.9
3	%	23.3	24.4	14.5	23.8	26.9	22.6
4	%	65.0	66.7	79.0	64.3	69.2	69.3
Principal Diagnosis (MCD)							
Duodenum tumour	%	5.0	2.2	11.3	4.8	6.4	6.3
Extrahepatic bile ducts or ampulla of Vater tumour	%	16.7	13.3	11.3	14.3	28.2	17.8
Pancreas tumour	%	78.3	84.4	77.4	81.0	65.4	76.0
Length of stay	Median	25.5	24.0	21.5	25.0	22.5	23.0
In-hospital + approximate 30-day mortality	%	10.0	6.7	11.3	7.1	6.4	8.4
2-year mortality	**%**	**55.0**	**60.0**	**59.7**	**45.2**	**42.3**	**51.9**
Stage							
0-I	%	28.9	13.8	26.7	4.2	25.5	22.2
II	%	51.1	65.5	57.8	62.5	56.4	57.6
III	%	13.3	3.4	11.1	16.7	1.8	8.6
IV	%	6.7	17.2	4.4	16.7	16.4	11.6

RELATION BETWEEN SURGEON VOLUME AND 2-YEAR MORTALITY RATE

Because of lack of an international cut off for surgeon volume in the literature, we used the volume threshold for oesophagectomies (≥ 6/year) which was identified in the previous section (see page 57). When this volume threshold is applied, low volume surgeons (1-5/year) have 58.1% 2-year mortality rate while high volume (≥ 6/year) surgeons have a 43.3% mortality rate (Table 5.30). Survival curve are presented in Figure 5.15.

Table 5.30: Pancreatic cancer surgery: 2-year mortality per surgeon based on international volume thresholds

Cut off		Surgeon volume category	Number surgeons	Number cases	2-year mortality	
					Nbr deaths	%
Cut off for oesophagectomy because of lack of cut off for pancreatectomy	6/year	1-5/year	100	167	97	58.1
		≥ 6 /year	9	120	52	43.3

Figure 5.15 Pancreatic cancer surgery: survival curve by surgeon volume (threshold 6/year)

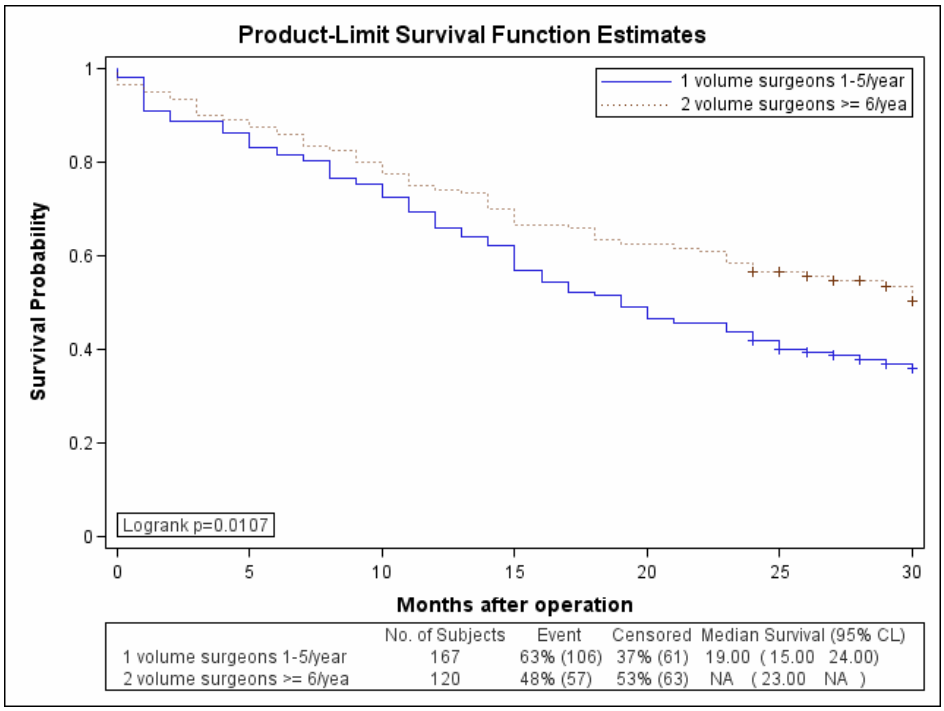


Table 5.31 shows the results from logistic regression (with GEE adjustment for clustering of patients within centres). Patients operated by surgeons performing less than 6 interventions per year have a 50% higher odds of mortality than patients operated by surgeons performing at least 6 interventions per year, taking into account patient and tumour characteristics: odds ratio and 95% CI 1.51 (1.06, 2.16).

Table 5.31: Pancreatic cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality

Model without adjustment for case mix				
Factor		OR	95%CI	
Surgeon Volume	1-5/year	1.80	1.06	3.07
	≥ 6 /year	1.00	1.00	1.00
Model with adjustment for patient characteristics (MCD data)				
Surgeon Volume	1-5/year	1.54	1.02	2.34
	≥ 6 /year	1.00	1.00	1.00
Sex	male vs female	1.30	0.83	2.05
Age	increase of 1 yr	1.03	1.00	1.06
Principal Diagnosis	Duodenum vs Pancreas	1.19	0.38	3.75
	Extrahepatic bile ducts or Ampulla of Vater versus Pancreas	0.42	0.22	0.79
Charlson score	increase of 1	1.07	0.79	1.46
Model with adjustment for patient and tumour characteristics (MCD and BCR data)				
Surgeon Volume	1-5/year	1.51	1.06	2.16
	≥ 6 /year	1.00	1.00	1.00
Sex	male vs female	1.46	0.90	2.36
Age	increase of 1 yr	1.04	1.00	1.08
Principal Diagnosis	Duodenum vs Pancreas	1.09	0.34	3.43
	Extrahepatic bile ducts or Ampulla of Vater versus Pancreas	0.42	0.19	0.94
Charlson score	increase of 1	1.07	0.78	1.46
Stage	unknown	1.20	0.52	2.74
	IV	2.06	0.60	7.08
	III	4.55	1.42	14.62
	II	2.35	1.12	4.94
	0-I	1.00	1.00	1.00

5.3.7 Discussion

5.3.7.1 External validation of the definition of the procedure

According to the Belgian Cancer Registry, there were 1 080 new cases of pancreas cancer in Belgium in 2004. See Table 5.1 (page 45). In the 2004 Minimal Clinical Data (MCD), we identified 1 526 patients with the principal diagnosis of malignant neoplasm of pancreas (see Table 5.19, page 66). The higher number of patients in the MCD is explainable by the fact that these hospital stays are not restricted to patients whose cancer started in 2004.

The study by Topal et al. supplies us with reference material.[156] Topal and colleagues obtained minimal clinical data (MCD) on all pancreaticoduodenectomies (PDs) performed in Belgium between 2000 and 2004. A total of 1 842 PDs were registered based on ICD-9-CM procedure code 527 and/or NIHDI billing code 242023. When we compare these selection criteria with ours (see Table 5.18), we notice that Topal et al. focused on only one operative technique, i.e. the pancreaticoduodenectomy (the Whipple procedure), which is the surgical procedure to remove the head of the pancreas. On the other hand, however, they did not select on principal diagnosis and therefore included patients with peripancreatic cancer or even with a disorder that is not cancer related such as chronic pancreatitis. Considering these differences in patient selection, it seems justifiable that Topal's absolute number of patients (1 842 over five years) is slightly higher than ours (320 in one year, see Table 5.19).

5.3.7.2 *Summarized results of literature review*

The systematic literature search identified 9 systematic reviews in which the volume outcome association (VOA) for pancreatic cancer surgery was studied.[1, 5, 59, 60, 62, 64, 69, 71, 72] These systematic reviews were based on 25 primary studies.[115-117, 129, 131, 133, 135, 141, 142, 155, 157-171] An additional search for more recent studies which were published in the period 2004-2009 resulted in 8 additional primary studies.[50, 153, 156, 172-176] This brings us to a total of 33 primary studies that analyzed the volume outcome association for pancreatic cancer surgery. For more insight in how these studies were retrieved see the Supplement.

On the basis of the systematic reviews it was concluded in Chapter 2 (see Table 2.2 on page 19) that there is evidence for an inverse relation between hospital volume and mortality for pancreatic cancer surgery. This means that mortality decreases when the number of procedures performed by a hospital increases. A similar relationship was found between surgeon volume and mortality.

With respect to pancreatic resection for malignant disease, Killeen et al. found that the number needed to treat for a high-volume provider to prevent one death is 10 to 15 patients.[69]

One minimal hospital volume threshold was retained from the literature search i.e. 11 pancreatic resections per annum. This thresholds correspond more or less with the thresholds in the systematic reviews, but, more importantly, it corresponds with the threshold issued by the US Agency for Healthcare Research and Quality (AHRQ) and the US Leapfrog Group.[29, 126]

Detailed results of the literature review are described in the evidence tables in Appendix 8.

5.3.7.3 *Comparative analysis of literature and Belgian data*

DEFINITION OF VOLUME

Figure 5.9 and Figure 5.10 illustrate that only 5 hospitals and 4 surgeons had an annual volume of at least 11 procedures which is the cut off defined by the US AHRQ and the US Leapfrog Group.[29, 126] Topal et al. found that only 7 out of 126 Belgian hospitals annually performed more than 10 pancreaticoduodenectomies between 2000 and 2004.[156] These high-volume hospitals accounted for 43.2% of patients. Topal et al. did not find apparent changes with regard to the annual number of PDs per hospital over the five years.

OUTCOME

In this KCE study (see Table 5.25), in-hospital and approximate 30-day mortality is 8.3% after pancreatectomy for peripancreatic cancer. Table 5.32 compares this Belgian outcome data with those published in the studies that were selected. Most authors used the Whipple procedure as only selection criteria.[117, 156, 162, 167] Only a few used a patient selection which was similar to ours i.e. pancreatectomy together with the diagnosis of (peri)pancreatic cancer.[155, 175] Regardless of patient selection, short-term mortality rate always fluctuated between 8 and 10%. None of these studies provided information on one- or two-year survival.

Table 5.32: Pancreatic cancer surgery: Mortality: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Topal 2007 [156]	Nordback 2002 [167]	Ho 2003 [117]	Gouma 2000 [162]	Glasgow 1996 [155]
Country Study period	Belgium 2004	Belgium 2000-2004	Finland 1990-1994	USA 1988-1998	NL 1994-1998	USA 1990-1994
Patient selection	Peripan-creatic cancer + surgery	Pancrea-ticoduo-denec-tomy	Pancrea-ticoduo-denec-tomy	Pancrea-ticoduo-denec-tomy	Pancrea-ticoduo-denec-tomy	Peripan-creatic cancer + surgery
In-hospital or 30-day mortality	8.3%	8.4%	10.6%	9.5%	10.1%	9.9%

PATIENT CASE MIX

Only the study by Glasgow et al. allowed a proper comparison of patients since their patient selection was similar to ours.[155] As shown in Table 5.33, age and sex distributions are very similar. With respect to tumour location, Glasgow et al. counted less pancreas tumours. Tumour stage was not available in Glasgow's data.

Table 5.33: Pancreatic cancer surgery: Patient and tumour characteristics: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Glasgow 1996[155]
Country	Belgium	USA
Study period	2004	1990-1994
Age (median)	66 years	65 years
Male	55.0%	50.6%
Tumour location		
Duodenum	3.6%	8.2%
Extrahepatic bile ducts	6.5%	8.9%
Ampulla of Vater	14.1%	20.5%
Pancreas	75.8%	62.4%

Note that we found only two studies that applied tumour characteristics for risk adjustment.[153, 155] Topal et al. did not apply any risk adjustment whatsoever.[156] The following characteristics were additionally used for risk adjustment in other studies but could not be applied in the KCE study because they are not available in MCD or BCR data:

- race;[115, 131, 153, 155]
- median household income;[115, 153]
- college education;[153]
- acuity of index admission i.e. elective, urgent or emergent;[115, 131, 153, 155]
- payer source i.e. government versus private;[155]
- type of surgical resection i.e. total or partial pancreatectomy.[155, 175]

With respect to adjustment for use of adjuvant treatment Birkmeyer et al. found that patients at low-volume hospitals versus high-volume hospitals were substantially less likely to receive adjuvant radiation (26% versus 52%) or chemotherapy (21% versus 36%) for pancreatic cancer. Nevertheless, adjusting for differences in adjuvant therapy only had a negligible effect in attenuating differences in 5-year survival.[153]

VOLUME OUTCOME ASSOCIATION

The present study suggests that patients treated in hospital performing less than 11 interventions per year (the AHRQ cut off) have a higher 2-year mortality, taking into account patient and tumour characteristics, than patients operating in hospitals performing at least 11 interventions per year (56.1% vs 48.2%, OR and 95% CI 1.25 (0.83, 1.89). This association did not reach statistical significance.

With respect to the relationship between surgeon volume and 2-year mortality, our study showed that patients operated by surgeons performing less than 6 interventions per year (the AHRQ cut off) have a higher risk of mortality, taking into account patient and tumour characteristics (OR 1.51 (95% CI 1.06, 2.16)).

These results are a confirmation of what we found in the literature review.

Key points on volume outcome association for peripancreatic cancer surgery

- **A total of 1 842 patients were hospitalized in 2004 (retrieved in MCD 2004) for a diagnosis of malignant neoplasm of pancreas (or peripancreas). 17% of these patients underwent a pancreatectomy.**
- **The population studied consisted of those 301 patients with pancreatectomy and whose data could be linked to IMA databases.**
- **In 80% of the cases, information on tumour could be retrieved in the BCR database. Data on stage was available for 67% of stays. There was a good agreement between the diagnosis in MCD and the tumour location in BCR.**
- **These interventions were performed in 74 centres by 112 surgeons. Five hospitals and four surgeons had a volume higher than 11 interventions per year (the US AHRQ and Leapfrog cut off).**
- **Two-year mortality was 53%. Regression models were fitted to assess the association between hospital or surgeon volume and outcome. The following factors were taken into account in all analyses: sex, age, principal diagnosis (pancreas or peripancreas), Charlson score (co morbidity) and tumour stage.**
- **Based on the study of systematic reviews, it was concluded that there is evidence for an inverse relation between hospital volume and mortality for pancreatic cancer surgery. One minimal hospital volume threshold was retained from the literature search i.e. 11 pancreatic resections per annum.**
- **The present study suggests that patients treated in hospital performing less than 11 interventions per year (the AHRQ cut off) have a higher 2-year mortality than patients operating in hospitals performing at least 11 interventions per year (56.1% vs 48.2%, OR and 95% CI 1.25 (0.83, 1.89).**
- **The literature review also concluded that there is inverse relationship between surgeon volume and mortality.**
- **Our findings were in agreement with literature: there was a statistically significant inverse association between the volume of surgeons and the 2-year mortality: 58% for surgeons performing less than 6 interventions per year and 43% for surgeons performing at least 6 interventions per year (OR and 95% CI 1.51 (1.06, 2.16).**
- **Data from two years are required to have more precise estimates.**
- **Because data were not retrieved or not available in the databases, the following characteristics could not be used for risk adjustment: use of (neo) adjuvant therapy (chemo- or radiation therapy), acuity of admission (elective versus urgent), intention of surgery (palliative versus curative), type of surgical resection (total or partial resection).**

5.4 COLON CANCER SURGERY

5.4.1 General description of procedure

Colectomy, which is a surgical procedure to remove the section of colon containing cancer and nearby lymph nodes, is the standard treatment for colon cancer and remains the only curative option.[120, 177]

5.4.2 Primary data selection in MCD 2004

First selection of data was based on ICD-9-CM diagnostic and procedure codes as proposed by Harmon et al. and Rabeneck et al., and on NIHDI procedure codes.[178, 179] Neoplasms of the rectosigmoid junction (i.e. the segment of the gut where the colon merges with the rectum) should also be included in the selection. Therefore, the Hartmann's procedure was added. This procedure refers to the anterior resection of the sigmoid and the rectum with closure of the distal resection margin and end colostomy. According to Vlayen et al., Hartmann's procedure is encoded with the non specific ICD-9-CM procedure code 45.75 (for left hemicolectomy) and the very specific NIHDI procedure code 244064.[180]

This selection resulted in a total of 20 264 stays. This selection is too broad since it also includes patients with a secondary diagnosis of malignant neoplasm. This implies that patients with colon cancer as pre-existing condition are also included (see section 4.2.1, page 39). The definition of the population studied will be defined in the next section.

Table 5.34: Colon cancer surgery: Primary data selection in MCD 2004

SELECTION 1	SELECTION 2	SELECTION 3
ICD-9-CM (principal or secondary) diagnosis code	**ICD-9-CM procedure code**	**NIHDI procedure code**
153x Malignant neoplasm of colon	457x Open and other partial excision of large intestine	243040 Total colectomy
1540 Malignant neoplasm of rectosigmoid junction	458x Total intra-abdominal colectomy	243062 Right or left hemicolectomy or segmentary colectomy or sigmoidectomy or partial resection of the rectum with restoration of continuity 243084 Segmentary colectomy with double colostomy 243106 Ileo-colo-recto plasty 244064 Hartmann's procedure
STAYS SELECTED = 14 383 stays	**STAYS SELECTED = 9 105 stays**	**STAYS SELECTED = 7 282 stays**
TOTAL STAYS SELECTED = 20 264 stays (selection 1 OR selection 2 OR selection 3)		

5.4.3 Definition of procedure

5.4.3.1 Primary hospital stays in Minimal Clinical Data

Approximately one third (i.e. 7 046 stays) of the 20 264 selected hospital stays had a principal diagnosis of malignant neoplasm of colon (80% i.e. 5 657 stays) or rectosigmoid (20% i.e. 1 389 stays).

During the analysis of the percentage of surgical procedures of these 7 046 stays, we noticed that 648 stays had only an ICD-9-CM procedure code without an accompanying NIHDI procedure code. This discrepancy probably finds his explanation in the fact that the selected ICD-9-CM procedure codes (see Table 5.34) are not specific enough for colon cancer surgery.

The author of KCE report 81 confirmed that he encountered the same problem during his study of rectum cancer.[180] Therefore, it was decided to limit the selection of patients to those with the diagnosis of colon or rectosigmoid cancer and at least one of the NIHDI procedure codes (see Table 5.34).

As shown in Table 5.35, 2 765 stays (39.2% of 7 046 stays) corresponding to 2 761 patients (47.9% of 5 756 patients) had a NIHDI procedure for colon cancer surgery as described above.

Table 5.35: Colon cancer surgery: Principal diagnosis and percentage of cancer surgery (per stay and per patient)

Principal diagnosis	**Total number hospital stays**	**Stays with colon cancer surgery** Nbr	%	**Total number patients**	**Patients with colon cancer surgery** Nbr	%
153 Malignant neoplasm of colon	5 657	2 623	46.4	4 641	2 619	56.4
1540 Malignant neoplasm of rectosigmoid junction	1 389	142	10.2	1 115	142	12.7
Total	7 046	2 765	39.2	5 756	2 761	48.0

The APR-DRGs of these 2 765 stays are described in Table 5.36.

To create a study population which is as homogenous as possible, we retained only the 2 730 stays (corresponding to 2 724 patients) that are included in the APR-DRG 221 "Major small & large bowel procedures".

Table 5.36: Colon cancer surgery: All APR-DRGs of stays with a principal diagnosis of a malignant neoplasm of colon or rectosigmoid junction AND with surgical resection of the colon

APR-DRG	Number	Percent
004-TRACHEOSTOMY EXCEPT FOR FACE, MOUTH & NECK DIAGNOSES / p3 - P	10	0.36
220-MAJOR STOMACH, ESOPHAGEAL & DUODENAL PROCEDURES / 6 - P	21	0.76
221-MAJOR SMALL & LARGE BOWEL PROCEDURES / 6 - P	**2730**	**98.73**
223-MINOR SMALL & LARGE BOWEL PROCEDURES / 6 - P	2	0.07
226-ANAL & STOMAL PROCEDURES / 6 - P	1	0.04
229-OTHER DIGESTIVE SYSTEM PROCEDURES / 6 - P	1	0.04
TOTAL	2765	100

Results of definition: 2 730 stays with colon cancer surgery were selected in the Minimal Clinical Data. These stays concerned 2 724 individual patients and 114 centres.

5.4.3.2 *Linkage with data from IMA and BCR*

The Minimal Clinical Data (MCD) data were linked with the Common Sickness Funds Agency (IMA) database and the Belgian Cancer Registry (BCR) database. Table 5.37 shows that 97.4% of 2 730 MCD stays could be linked with IMA data and 80.5% with BCR data (on the basis of ICD-10 codes C18 and C19). Information on tumour stage could be retrieved for 71.6% of 2 730 MCD stays.

Table 5.37: Colon cancer surgery: Percentage of linkage of MCD data with data from IMA and BCR

	Number	%
Number of stays in MCD selection	2 730	
Linkage with IMA	2 658	97.4
Linkage with BCR	2 198	80.5
Linkage with BCR and data about stage	1 955	71.6

Results of linkage: 2 198 patients with colon cancer surgery could be linked with Belgian Cancer Registry data. For 1 955 of these patients, the BCR contained information on the stage of the tumour.

5.4.3.3 *Patient and tumour characteristics*

Tumour location is defined according to the ICD-10 code:

- C18 for malignant neoplasm of colon
- C19 for malignant neoplasm of rectosigmoid junction.

Tumour histology is not used as a variable for risk adjustment and is therefore not analyzed in this section. For a detailed description of TNM classification and TNM stages for colon cancer, we refer to the Belgian national guidelines.[120, 177]

As shown in Table 5.38, mean age of patients was 72 years and 50% of them were men. Majority of patients had a colon cancer (98%). Forty per cent of patients had stage 2 colon cancer, 35% had stage 3 disease.

Table 5.38: Colon cancer surgery: Patient and tumour characteristics

		Number	%
Age (years)			
	mean/median	72.3/74	
	Std	11.2	
	Total	2730	
Sex			
	Male	1366	50.0
	Female	1364	50.0
	Total	2730	
Tumour location			
	Colon	2152	97.9
	Rectosigmoid junction	46	2.1
	Total	2198	
Tumour stage			
	I	244	12.5
	II	797	40.8
	III	685	35.0
	IV	229	11.7
	Total		

5.4.4 Definition of volume

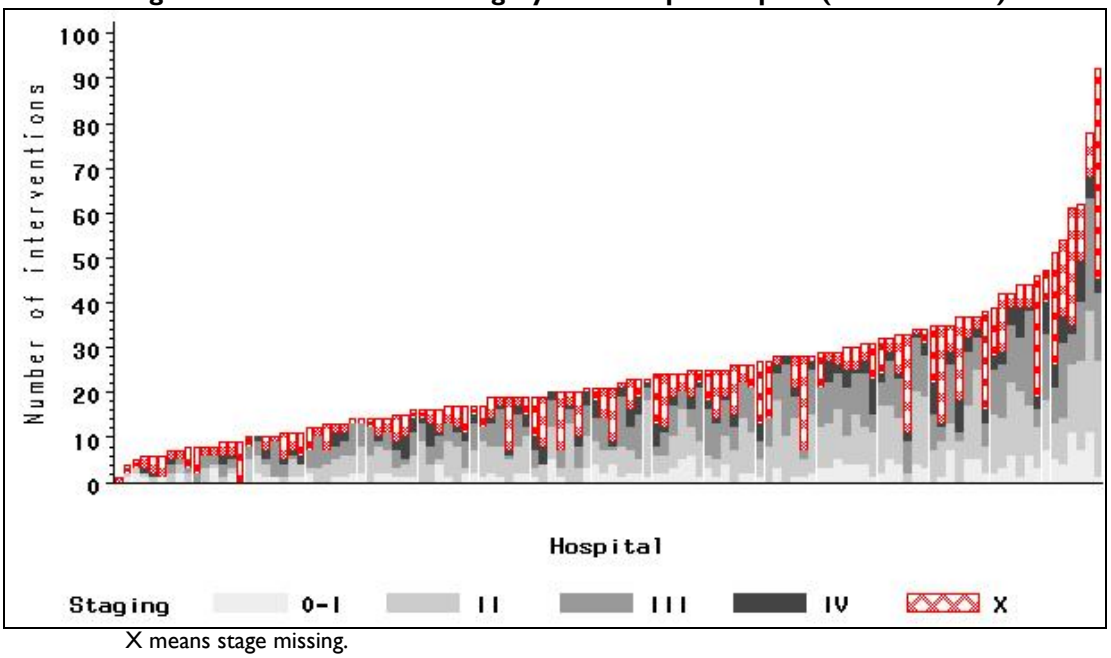
These 2 730 procedures were performed in 114 centres by 401 surgeons. Three quarters of the hospitals performed 30 procedures or less.

Table 5.39: Colon cancer surgery: Summary measures of volume per hospital and per surgeon

	Number	mean	min	25th Pctl	50th Pctl	75th Pctl	Maximum
Hospitals	114	23.9	1.0	14.0	21.0	30.0	92.0
Surgeons*	401	6.8	1.0	3.0	5.0	9.0	52.0

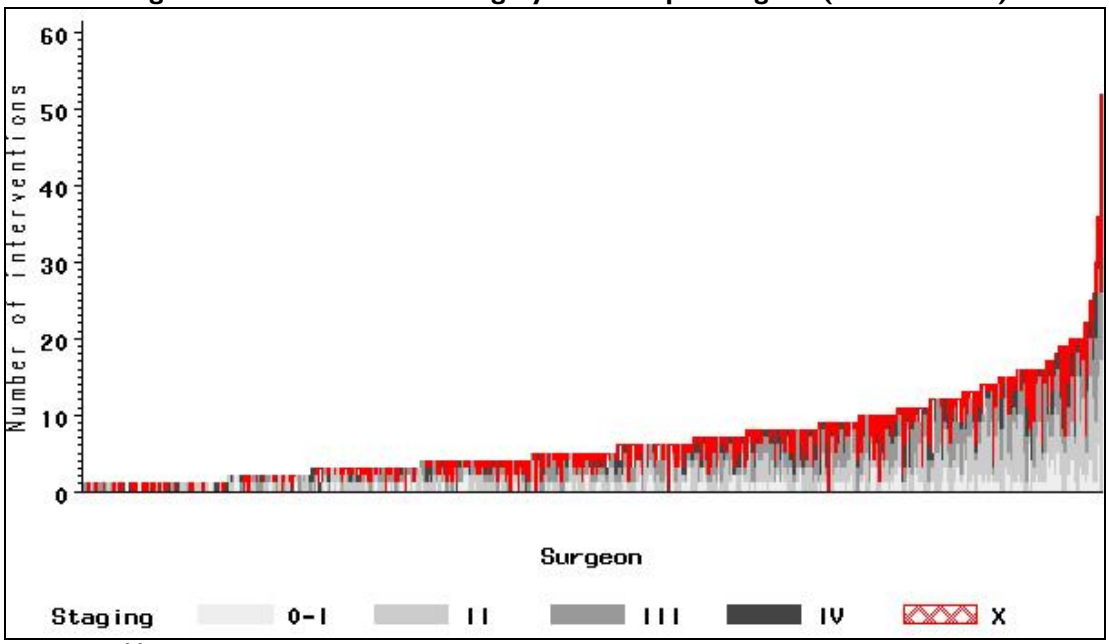
* For 10 procedures (0.4%) the information on surgeon ID was missing.

Figure 5.16: Colon cancer surgery: Volume per hospital (Number=114)



X means stage missing.

Figure 5.17: Colon cancer surgery: Volume per surgeon (Number=401)



X means stage missing.

5.4.5 Definition of outcomes

5.4.5.1 Mortality

Table 5.40 shows that in-hospital and approximate 30-day mortality is 5.6% after colon cancer surgery (148 deaths). Mortality 1 year after surgery is 18.0%. After 2 years it is 28.2%. Mean survival time (median not defined) was 27 months (not visible in Table 5.40).

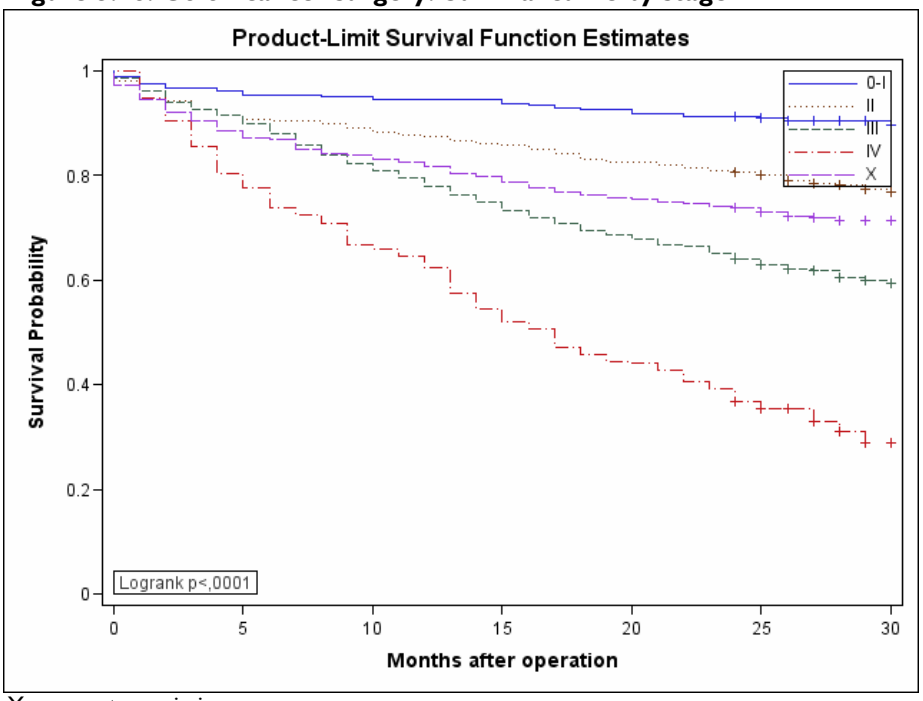
Table 5.40 also shows that 2-year mortality after surgery is higher with patients diagnosed with a cancer of the rectosigmoid junction (4310%) in comparison with patients suffering from a colon cancer (27.4%). These differences confirm the necessity to add the principal diagnosis to the risk adjustment.

Table 5.40: Colon Cancer surgery: Mortality results

	Number	Number deaths	% deaths	Survival
In-hospital and approximate 30-day mortality	2658	148	5.57	
1-year mortality	2658	479	18.02	81.98
2-year mortality	2658	750	28.22	71.78
2-year mortality in function of principal diagnosis:				
153 Malignant neoplasm of colon	2521	691	27.41	72.59
1540 Malignant neoplasm of rectosigmoid junction	137	59	43.07	56.93

Figure 5.18 shows how survival improves with decreasing disease stage. Patients with stage 3 and 4 disease have the worst survival curves while stage 1 and 2 patients have the best survival. Patients with unknown disease stage are in between these two groups.

Figure 5.18: Colon cancer surgery: Survival curve by stage



X means stage missing.

Because of lack of reliable data on complications, all analyses on the volume-outcome association are limited to 2-year mortality rate. This implies that additional outcome measures such as postoperative rate of stoma, infection and anastomotic leakage, which are listed in Table 2.1 (page 16), are not analyzed in this report.

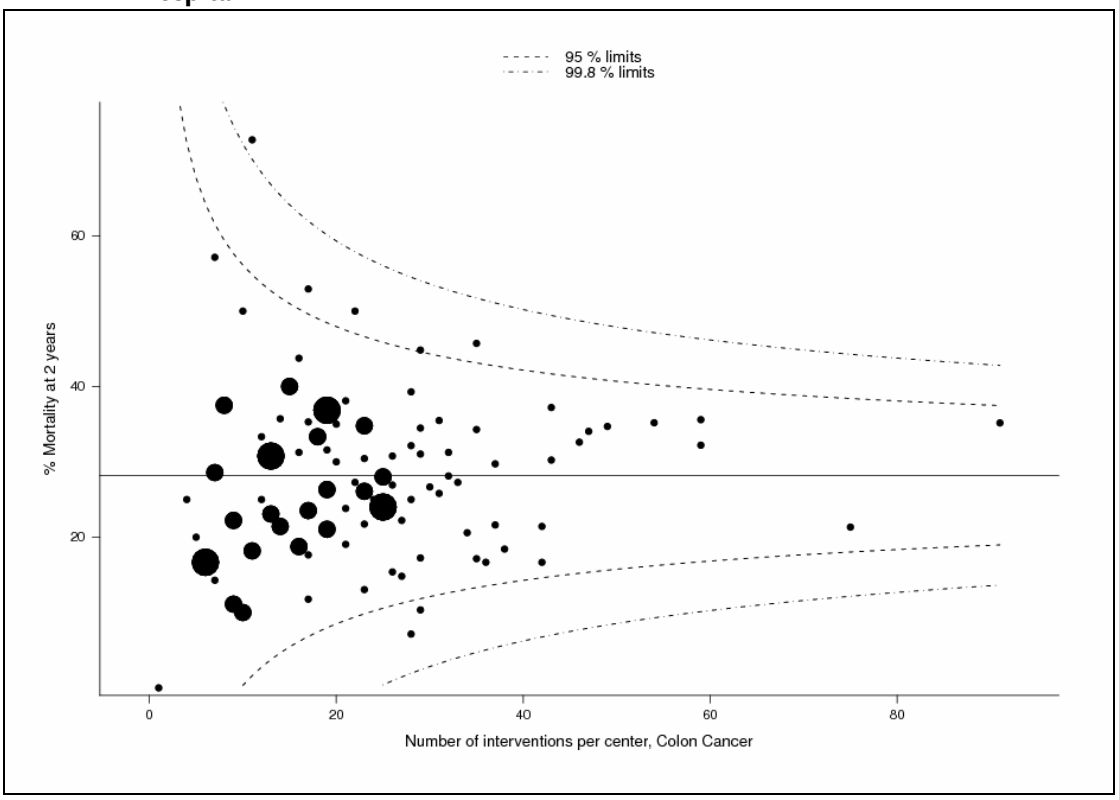
5.4.6 Volume outcome relationship

5.4.6.1 Analysis by hospital

2-YEAR MORTALITY RATE, BY HOSPITAL

Figure 5.19 presents the funnel plot of the number of colon cancer procedures per centre and the 2-year mortality (observed, i.e. without risk adjustment). The size of the point is proportional to the number of hospitals with the same volume and the same outcome. The horizontal line represents the overall 2-year mortality i.e. 28.2%. Only one small volume hospital is outlier above the 99.8% limits of variability.

Figure 5.19: Colon cancer surgery: Funnel plot of 2-year mortality rate, by hospital



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

As shown in Table 5.41, 2 year mortality is 24% in lowest volume hospitals (1-10 procedures/yr) and around 31% for centres performing at least 41 interventions. The percentage of missing stage data reaches 39% in lowest volume hospitals (compared to 26% in all hospitals), so comparison of stages between centres should be made cautiously.

Table 5.41: Colon cancer surgery: Differences in case-mix and outcomes by hospital volume

		Hospital volume					
		1-10/yr	11-20/yr	21-40/yr	41-60/yr	>60/yr	All
Number centres		19	35	48	8	4	114
Number stays		138	552	1318	366	284	2658
Gender	%male	56.5	47.6	50.1	50.8	51.1	50.1
	Mean	70.9	72.5	72.3	72.8	71.4	72.3
Age	Std	10.8	10.9	11.2	10.9	12.3	11.2
Charlson ≥ 3	%	6.5	12.1	11.5	8.5	12.0	11.0
	Mean	0.8	0.9	0.9	0.8	1.0	0.9
Charlson	Std	1.2	1.3	1.4	1.2	1.5	1.4
Severity index (APR-DRG)							
1	%	6.5	10.3	10.1	9.0	5.6	9.3
2	%	39.9	29.7	35.2	28.1	44.4	34.3
3	%	38.4	41.5	40.6	44.5	33.8	40.5
4	%	15.2	18.5	14.1	18.3	16.2	15.9
Principal Diagnosis (MCD)							
Colon tumour	%	92.8	94.7	95.1	95.6	94.0	94.8
Rectosigmoid junction tumour	%	7.2	5.3	4.9	4.4	6.0	5.2
Length of stay	Median	15.5	14.0	15.0	14.0	13.0	14.0
In-hospital + approximate 30-day mortality	%	4.3	4.2	4.6	4.4	3.2	4.3
2-year mortality	**%**	**23.9**	**29.9**	**26.7**	**30.6**	**31.0**	**28.2**
Missing stage data	%	39.1	27.7	23.6	26.8	30.6	26.4
Stage							
I	%	19.0	10.3	12.1	13.1	15.2	12.5
II	%	33.3	42.4	41.4	35.4	44.7	40.8
III	%	36.9	34.3	35.3	38.1	30.5	35.0
IV	%	10.7	13.0	11.2	13.4	9.6	11.7

RELATION BETWEEN HOSPITAL VOLUME AND 2-YEAR MORTALITY RATE

None of the models in Table 5.42 showed a relationship between hospital volume and 2-year mortality rate. All OR estimates are below and close to 1, except in the first volume category, where OR is 0.64 (0.38, 1.07). Survival rate is lower in patients with a colon cancer when compared with those with cancer of the rectosigmoid junction (OR = 0.44, 95%CI 0.30-0.65). Disease stage is again an important predictor of 2-year mortality. Patients with stage 2, 3 and 4 disease have significant higher 2-year mortality than those with stage 1 disease.

Table 5.42: Colon cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality

Model without adjustment for case mix				
Factor		OR	95%CI	
Hospital Volume	1-10/year	0.70	0.44	1.11
	11-20/year	0.95	0.69	1.31
	21-40/year	0.81	0.60	1.10
	41-60/year	0.98	0.69	1.39
	> 60/year	1.00	1.00	1.00
Model with adjustment for patient characteristics (MCD data)				
Hospital Volume	1-10/year	0.71	0.44	1.15
	11-20/year	0.93	0.66	1.31
	21-40/year	0.79	0.57	1.09
	41-60/year	0.97	0.68	1.38
	> 60/year	1.00	1.00	1.00
Sex	male vs female	1.13	0.95	1.35
Age	increase of 1 yr	1.05	1.04	1.06
Principal Diagnosis	Malignant neoplasm of colon vs of rectosigmoid junction	0.46	0.32	0.67
Charlson score	increase of 1	1.28	1.17	1.41
Model with adjustment for patient and tumour characteristics (MCD and BCR data)				
Hospital Volume	1-10/year	0.64	0.38	1.07
	11-20/year	0.79	0.53	1.17
	21-40/year	0.68	0.47	0.97
	41-60/year	0.80	0.54	1.18
	> 60/year	1.00	1.00	1.00
Sex	male vs female	1.20	0.99	1.44
Age	increase of 1 yr	1.06	1.05	1.07
Principal Diagnosis	Malignant neoplasm of colon vs of rectosigmoid junction	0.44	0.30	0.65
Charlson score	increase of 1	1.35	1.21	1.50
Stage	unknown	3.90	2.24	6.80
	IV	27.95	13.98	55.84
	III	6.99	4.08	11.97
	II	2.52	1.39	4.55
	I	1.00	1.00	1.00

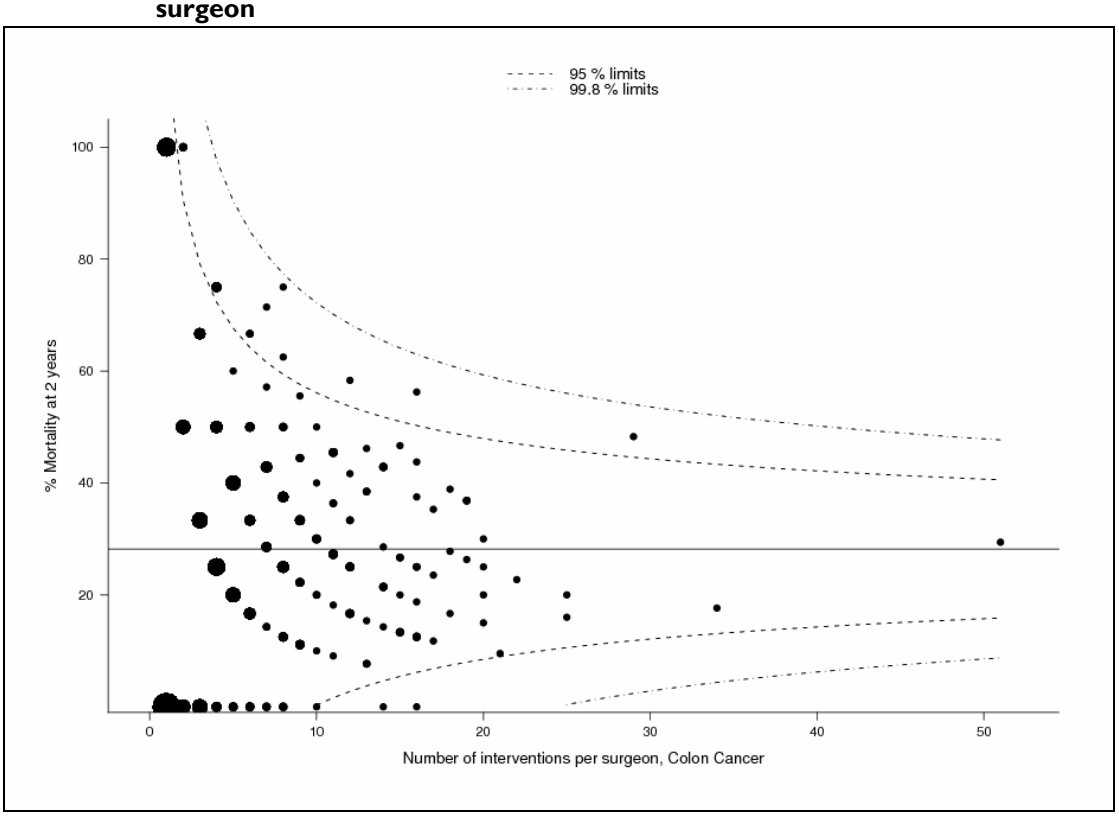
Sensitivity analyses with the log of volume as a predictor variable confirmed those results (positive coefficient of volume effect).

5.4.6.2 Analysis by surgeon

2-YEAR MORTALITY RATE, BY SURGEON

Figure 5.20 presents the funnel plot of the number of colon cancer procedures per surgeon and the 2-year mortality (observed, i.e. without risk adjustment). The size of the point is proportional to the number of surgeons with the same volume and the same outcome. There are no surgeons outside the 99.8% limits of variability.

Figure 5.20: Colon cancer surgery: Funnel plot of 2-year mortality rate, by surgeon



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE SURGEON

As shown in Table 5.43, 2-year mortality is 30.8% in lowest volume surgeons (1-5 procedures/yr) and 24.6% in surgeons which perform more than 20 colon cancer procedures per year. Patients who are treated by low-volume surgeons seem to have more co morbidities (i.e. more patients with Charlson score ≥ 3). However, distribution of the tumour stages among the surgeon categories does not seem to be related to the surgeon volume.

Table 5.43: Colon cancer surgery: Differences in case-mix and outcomes by surgeon volume

		Surgeon Volume					
		1-5/yr	**6-10/yr**	**11-15/yr**	**16-20/yr**	**> 20 /yr**	**All**
Number surgeons		209	110	47	27	7	400
Number stays		581	825	586	449	207	2648
Gender	% male	52.5	48.2	48.8	50.6	53.1	50.1
Age	Mean	72.0	72.5	72.7	72.0	71.4	72.3
	Std	12.3	10.7	10.8	10.9	11.5	11.2
Charlson ≥ 3	%	12.2	12.0	9.7	10.0	8.2	10.9
Charlson	Mean	1.0	0.9	0.7	0.8	0.9	0.9
	Std	1.4	1.4	1.3	1.3	1.3	1.4
Severity index (APR-DRG)							
1	%	9.0	9.9	10.4	7.6	9.2	9.4
2	%	30.8	32.2	36.9	35.6	43.5	34.4
3	%	39.6	42.8	37.0	43.9	38.2	40.6
4	%	20.7	15.0	15.7	12.9	9.2	15.6
Principal diagnosis (MCD)							
colon	%	93.6	94.1	97.1	96.0	92.8	94.9
rectosigmoid junction	%	6.4	5.9	2.9	4.0	7.2	5.1
LOS	Mean	20.6	19.6	18.6	17.0	14.8	18.8
	Median	16.0	15.0	14.0	13.0	11.0	14.0
	Std	14.9	15.0	14.1	11.8	11.4	14.1
In-hospital + approximate 30-day mortality	%	7.2	6.3	4.3	5.1	1.9	5.5
2-year mortality	**%**	**30.8**	**29.6**	**26.6**	**26.1**	**24.6**	**28.2**
Missing stage BCR	%	31.2	28.5	21.7	20.7	30.0	26.4
Stage							
I	%	11.8	12.7	12.2	14.3	10.3	12.5
II	%	40.3	37.6	43.6	41.6	43.4	40.7
III	%	36.0	36.3	33.6	32.9	37.2	35.0
IV	%	12.0	13.4	10.7	11.2	9.0	11.7

RELATION BETWEEN SURGEON VOLUME AND 2-YEAR MORTALITY RATE

Table 5.44 shows the results from logistic regression (with GEE adjustment for clustering of patients within centres). All three models only suggest that patients treated by low-volume surgeons (1-5 procedures/yr) have a higher 2 year mortality rate than patients operated by high-volume surgeons (>20/yr), without reaching statistical significance.

Table 5.44: Colon cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality (surgeon volume)

Model without adjustment for case mix				
Factor		OR	95%CI	
Surgeon volume	1-5/year	1.38	0.82	2.33
	6-10/year	1.30	0.78	2.18
	11-15/year	1.13	0.67	1.91
	16-20/year	1.08	0.63	1.85
	> 20 /year	1.00	1.00	1.00
Model with adjustment for patient characteristics (MCD data)				
Surgeon volume	1-5/year	1.34	0.78	2.31
	6-10/year	1.27	0.74	2.16
	11-15/year	1.14	0.66	1.95
	16-20/year	1.08	0.62	1.89
	> 20 /year	1.00	1.00	1.00
sex	male vs female	1.14	0.96	1.36
age	increase of 1 y.	1.05	1.04	1.06
Principal diagnosis	Malignant neoplasm of colon vs of rectosigmoid junction	0.46	0.32	0.67
Charlson score	increase of 1	1.27	1.16	1.40
Model with adjustment for patient and tumour characteristics (MCD and BCR data)				
Surgeon volume	1-5/year	1.36	0.84	2.20
	6-10/year	1.24	0.79	1.95
	11-15/year	1.17	0.71	1.92
	16-20/year	1.11	0.68	1.81
	> 20 /year	1.00	1.00	1.00
sex	male vs female	1.21	1.00	1.46
age	increase of 1 y.	1.06	1.05	1.07
Principal diagnosis	Malignant neoplasm of colon vs of rectosigmoid junction	0.44	0.30	0.65
Charlson score	increase of 1	1.34	1.20	1.49
Stage	unknown	3.73	2.11	6.62
	IV	27.17	13.55	54.50
	III	6.89	4.02	11.80
	II	2.51	1.38	4.54
	I	1.00	1.00	1.00

Table 5.45 shows the effects of logistic regression when volume of centres and surgeons are taken into account simultaneously. This analysis confirmed the effects observed when volume of centre and surgeon were analyzed separately. The effect of volume of surgeon was consistent (inverse relationship) in low volume centres and high volume centres (no interaction between the two effects, data not shown).

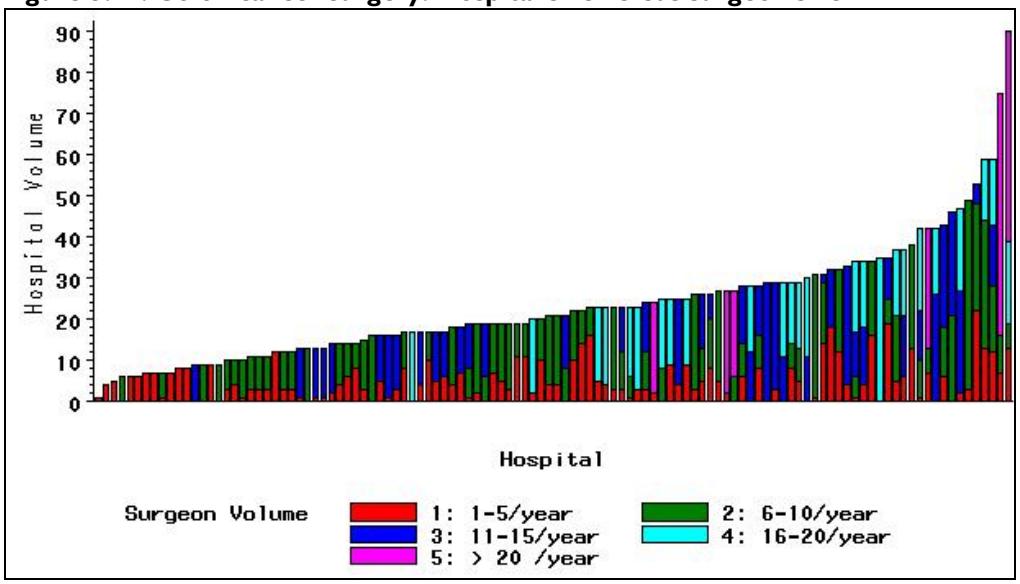
Table 5.45: Colon cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality, surgeon volume and hospital volume

Parameter	Estimate	Standard Error	95% Confidence Limits	
Log volume of surgeon	-0.1491	0.0735	-0.2933	-0.0050
Log volume of hospital	0.1425	0.1064	-0.0661	0.3510

After adjustment for gender, age, Charlson score, principal diagnosis and stage

Figure 5.21 shows that high volume surgeons (in pink) tend to operate in high volume centres, while low volume surgeons (in red) tend to operate in all centres.

Figure 5.21: Colon cancer surgery: Hospital size versus surgeon size



5.4.7 Discussion

5.4.7.1 External validation of the definition of the procedure

The Belgian Cancer Registry registered 5 482 new cases of colon and rectosigmoid junction cancer in Belgium in 2004. See Table 5.1 (page 45). This number corresponds well to the 5 756 patients with the principal diagnosis of these cancers which were identified in the 2004 Minimal Clinical Data (see Table 5.35, page 84).

5.4.7.2 Summarized results of literature review

The systematic literature search identified 9 systematic reviews in which the volume outcome association (VOA) for colon cancer surgery was studied.[1, 5, 59, 60, 62, 64, 69, 73-75] These systematic reviews were based on 43 primary studies.[99, 115, 129, 133-135, 141, 142, 178, 179, 181-212, 213] An additional search for more recent studies which were published in the period 2004-2009 resulted in 5 additional primary studies.[153, 214-217] This brings us to a total of 48 primary studies that analyzed the volume outcome association for colon cancer surgery. For more insight in how these studies were retrieved see the Supplement.

On the basis of the systematic reviews it was concluded in Chapter 2 (see Table 2.2 on page 19) that there is evidence for an inverse relation between hospital volume and mortality for colon cancer surgery.

This means that mortality decreases when the number of procedures performed by a hospital increases. A similar relationship was found between surgeon volume and mortality.

However, the effect of volume on outcome for this procedure, which is more common and often a less complex surgical procedure, is small when compared to other, more complex procedures such as pancreatic or oesophageal resection. Killeen et al. calculated that the magnitude of the volume effect on mortality was variable and small (approximately 1–2 per cent). They translated this into a number needed to treat of 50–100 patients.[69]

With respect to a cut off point to define low and high volume, studies differed widely. Therefore, it was impossible to define one minimal hospital volume threshold. For example, Killeen et al., found that the definition of low volume ranged from one to fewer than 12 procedures for low-volume surgeons and from one to fewer than 84 operations per year for low-volume hospitals.[69]

Detailed results of the literature review are described in the evidence tables in Appendix 8.

5.4.7.3 Comparative analysis of literature and Belgian data

DEFINITION OF VOLUME

As mentioned before, it was impossible to define a cut off point for hospital or surgeon volume on the basis of the literature review.

OUTCOME

With respect to the comparison of Belgian outcome data with those published in the studies that were selected (see 5.4.7.2), it was difficult to find studies that made the same patient selection as we did.

- Many authors included rectum cancers in their selection.[185, 201] However, according to Dimick et al. mortality rates vary depending on the location of the tumour and are lowest for tumour of the rectum (2.4%) and highest for tumours of the left colon (4.5%).[185] Therefore, it seems not right to compare the mortality rate of colorectal cancer studies to our colon cancer study.
- Ko et al., on the other hand, only selected cases with the diagnosis of colon cancer and excluded rectal and rectosigmoid cancers.[198] As was shown in Table 5.40, patients with cancer of the rectosigmoid junction had worse survival (57.0%) than those with colon cancer (72.6%). Again a reason not to compare Ko's mortality rate with ours.
- Schrag et al. defined a similar population as we did i.e. colon and rectosigmoid cancer patients with resection.[209, 210] However, they restricted their cohort to patients with a histological diagnosis of adenocarcinoma. In addition, patients were all 65 years or older since it concerned Medicare-enrolled patients. Schrag et al. found a 30-day mortality rate of 4.6%, a 2-year survival rate of 66.9% and a 5-year survival rate of 47.3%.[209, 210] Marusch and colleagues also specifically investigated only patients with carcinoma of the colon.[200] They found an overall 30-day mortality rate of 4.9%.

Table 5.46: Colon cancer surgery: Mortality: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Schrag 2000 [209] & Schrag 2003 [210]	Marusch 2001 [200]
Country & Study period	Belgium 2004	USA 1991-1996	Germany 1999
30-day mortality	5.6%	4.6%	4.9%
1-year survival	82.0%		
2-year survival	71.8%	66.9%	

PATIENT CASE MIX

The studies by Schrag and Marusch allowed a proper comparison of patients since their patient selection was similar to ours.[200, 209, 210] See Table 5.47.

Table 5.47: Colon cancer surgery: Patient and tumour characteristics: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Schrag 2000[209]	Marusch 2001[200]
Country	Belgium	USA	Germany
Study period	2004	1991-1996	1999
Age (mean)	72.3 yr		68.5 yr
Male	50.0%	44.8%	51.6%
Tumour stage			
unknown		5.1%	2.1%
I	12.5%	19.2%	18.3%
II	40.8%	36.2%	31.0%
III	35.0%	25.6%	29.5%
IV	11.7%	13.9%	19.2%

We found several studies that applied tumour characteristics for risk adjustment.[153, 209, 210] The following characteristics were additionally used for risk adjustment in other studies but could not be applied in the KCE study because they are not available in MCD or BCR data:

- race;[115, 153, 185, 209, 210]
- median household income;[115, 153, 185, 209, 210]
- college education;[153]
- acuity of index admission i.e. elective, urgent or emergent;[115, 153, 185, 198, 209, 210]
- use of adjuvant therapy;[153, 209]
- urgency of surgery i.e. perforation or obstruction;[209, 210]
- type of colectomy i.e. right, left, sigmoid, abdominoperineal or transverse.[185]

With respect to adjustment for use of adjuvant treatment Birkmeyer and Schrag found that patients at low-volume hospitals were less likely to receive adjuvant chemotherapy for colon cancer. Nevertheless, adjusting for differences in adjuvant therapy only had a negligible effect in attenuating differences in 5-year survival.[153, 209]

VOLUME OUTCOME ASSOCIATION

The present study found no association between hospital volume and 2-year mortality. Two-year mortality was unexpectedly low in very low volume centres (23.9% in very low volume centres (1-10/year compared to average of 28.2), but these are also the centres with the highest percentage of missing stage data (40% compared to average 26%), This result contrasts with the findings from the literature review which concluded that there is evidence for an inverse relationship between hospital volume and mortality for colon cancer surgery.

A possible explanation for the fact that low-volume hospitals have such low mortality could be that these patients were treated by high-volume surgeons. This hypothesis was tested, however, and rejected since Figure 5.21 showed that patients at low-volume hospitals were mainly operated on by low-volume surgeons.

With respect to surgeon volume, the study suggests an inverse association with 2-year mortality, although it did not reach statistical significance. Two-year mortality decreased from 30.8% for small volume surgeons (<6/year) and 24.6% for high volume surgeons (>20/year). As other authors pointed out, this effect may also result from the skill of the individual surgeon. Porter et al. found that cancer-specific survival is improved with both colorectal surgical subspecialty training and a higher frequency of rectal cancer surgery. Therefore, they recommend that the surgical treatment of rectal cancer patients should rely exclusively on surgeons with such training or surgeons with more experience.[218] McArdle and colleagues go even further when they support that surgical specialization is the primary determinant of outcome for colorectal cancer.[201]

Key points on volume outcome association for colon cancer surgery

- **A total of 5 756 patients were hospitalized in 2004 (retrieved in MCD 2004) with a diagnosis of malignant neoplasm of colon (including rectosigmoid junction). 48% of these patients underwent a colectomy.**
- **The population studied consisted of those 2 658 patients with colon cancer surgery and whose data could be linked to IMA databases.**
- **In 81% of the cases, information on tumour could be retrieved in the BCR database. Data on stage was available for 72% of stays. In low volume centres (<10/year) this was only 61%.**
- **These interventions were performed in 114 centres by 401 surgeons. Half of the surgeons performed less than 6 procedures per year. There is no international cut off of volume for this procedure.**
- **Two-year mortality was 28%. Regression models were fitted to assess the association between hospital or surgeon volume on this outcome. The following factors were taken into account in all analyses: sex; age, principal diagnosis (colon or rectosigmoid junction), Charlson score (co morbidity) and tumour stage.**
- **Based on the study of systematic reviews, it was concluded that there is evidence for an inverse relation between hospital volume and mortality for colon cancer surgery. However, the effect of volume on outcome for this more common procedure is small when compared to pancreatic or oesophageal resection.**
- **The present study found no association between hospital volume and 2-year mortality. Two-year mortality was unexpectedly low in very low volume centres (23.9% in very low volume centres (1-10/year compared to average of 28.2%), but these are also the centres with the highest percentage of missing stage data (40% compared to average 26%).**
- **The literature review also concluded that there is inverse relationship between surgeon volume and mortality.**
- **This inverse relationship between volume of surgeons and 2-year mortality was suggested by Belgian data: approximately 30% for small volume surgeons (1-10/year), and 25% for high volume surgeons (> 20 years). Other authors suggested that surgical specialization is an important additional determinant of outcome for colorectal cancer.**
- **Because data were not retrieved or not available in the databases, the following characteristics could not be used for risk adjustment: use of (neo) adjuvant therapy (chemo- or radiation therapy), acuity of admission (elective versus urgent), intention of surgery (palliative versus curative), type of surgical resection (total or partial resection).**

5.5 BREAST CANCER SURGERY

5.5.1 General description of procedure

As reported by the Belgian Cancer Registry, breast cancer is the most frequent cancer in females (35%) and the leading cause of death by cancer in females (20.6% of all cancer deaths).[108] Surgery is still the first-line treatment for localized breast cancer and consists of two kinds i.e. breast-conserving therapy and breast-ablative therapy. Both procedures carry a low risk for morbidity and mortality.[48, 219]

5.5.2 Primary data selection in MCD 2004

First selection of data was based on ICD-9-CM diagnostic and procedure codes as proposed by Roohan et al. and Guller et al., and on NIHDI procedure codes.[219, 220]

This selection resulted in a total of 24 871 stays but is too broad since it also includes patients with a secondary diagnosis of malignant neoplasm. This implies that patients with breast cancer as pre-existing condition are also included (see section 4.2.1, page 39). The definition of the population studied will be defined in the next section.

Table 5.48: Breast cancer surgery: Primary data selection in MCD 2004

SELECTION 1	SELECTION 2	SELECTION 3
ICD-9-CM (principal or secondary) diagnosis code	**ICD-9-CM procedure code**	**NIHDI procedure code**
174x Malignant neoplasm of female breast	8520 Excision or destruction of breast tissue, not otherwise specified 8521 Local excision of lesion of breast 8522 Resection of quadrant of breast 8523 Subtotal mastectomy 854x Mastectomy	226962 Urban procedure (i.e. extended radical mastectomy). 226984 Halsted procedure (radical mastectomy) or Pattey procedure (modified radical mastectomy) with ex tempore histological analysis. 227006 Halsted procedure (radical mastectomy) or Pattey procedure (modified radical mastectomy). 227021 Excision of a tumour out the weak tissues above the muscle fascia but with complete resection of the organ in which the tumour is located. 227043 Removal of a tumour or a cyst from the breast. 227065 Partial mastectomy or lumpectomy with axillary clearance.
STAYS SELECTED = 21 555 stays	**STAYS SELECTED = 12 935 stays**	**STAYS SELECTED = 12 550 stays**
TOTAL STAYS SELECTED = 24 871 stays (selection 1 OR selection 2 OR selection 3)		

5.5.3 Definition of procedure

5.5.3.1 Primary hospital stays in Minimal Clinical Data

It was only after execution of the primary selection that we realized that carcinoma in situ of the breast (ICD-9-CM code 233.0) should have been added as a primary selection criterion. It was too late, though, to retrieve all hospital stays with this diagnosis. Nevertheless, we retained the stays with carcinoma in situ during which breast surgery was performed. This implies that hospital stays with the same diagnosis but without surgery are not included in this analysis, which explains the blanks in Table 5.49.

11 585 of the 24 871 selected hospital stays had a principal diagnosis of malignant neoplasm of female breast. As shown in Table 5.49, 9 817 stays corresponding to 9 166 patients had a procedure for breast cancer surgery as described above.

Table 5.49: Breast cancer surgery: Principal diagnosis and percentage of cancer surgery (per stay and per patient)

Principal diagnosis	**Total number hospital stays**	**Stays with breast cancer surgery**		**Total number patients**	**Patients with breast cancer surgery**	
		Number	%		Number	%
174 Malignant neoplasm of female breast 2330 Carcinoma in situ of breast	11 585	9 215 602	79.54	10 048	8 625 541	85.84
Total		9 817			9 166	

The APR-DRGs of these 9 817 stays are described in Table 5.50.

To create a study population of patients which is as homogenous as possible, we retained only the 9 633 stays (corresponding to 8 992 patients) that are included in the APR-DRGs 362 "Mastectomy procedures" and 363 "Breast procedures except mastectomy".

Table 5.50: Breast cancer surgery: All APR-DRGs of stays with a principal diagnosis of malignant neoplasm of breast or carcinoma in situ AND with surgical resection of the breast

APR-DRG	Number	Percent
004-TRACHEOSTOMY EXCEPT FOR FACE, MOUTH & NECK DIAGNOSES / p3 - P	1	0.01
361-SKIN GRAFT & WOUND DEBRID EXC FOR SKIN ULCER & CELLULITIS / 9 - P	140	1.43
362-MASTECTOMY PROCEDURES / 9 - P	**5 291**	**53.90**
363-BREAST PROCEDURES EXCEPT MASTECTOMY / 9 - P	**4 342**	**44.23**
364-OTHER SKIN, SUBCUTANEOUS TISSUE & BREAST PROCEDURES / 9 - P	42	0.43
956-UNGROUPABLE / 0 - M	1	0.01
TOTAL	9 817	100.00

Results of definition: 9 633 stays with breast cancer surgery were selected in the Minimal Clinical Data. These stays concerned 8 992 individual patients and 114 centres.

5.5.3.2 *Linkage with data from IMA and BCR*

The Minimal Clinical Data (MCD) data were linked with the Common Sickness Funds Agency (IMA) database and the Belgian Cancer Registry (BCR) database. Table 5.51 shows that 98.1% of 9 633 MCD stays could be linked with IMA data and 87.4% with BCR data (on the basis of ICD-10 codes C50 and D05). Information on tumour stage could be retrieved for 79.3% of 9 633 MCD stays.

Table 5.51: Breast cancer surgery: Percentage of linkage of MCD data with data from IMA and BCR

	Number	%
Number of stays in MCD selection	9 633	
Linkage with IMA	9 450	98.10
Linkage with BCR	8 416	87.37
Linkage with BCR and data about stage	7 643	79.34

Results of linkage: 8 416 patients with breast cancer surgery could be linked with Belgian Cancer Registry data. For 7 643 of these patients, the BCR contained information on the stage of the tumour.

5.5.3.3 *Patient and tumour characteristics*

Tumour location is defined according to ICD-10 code:

- C50 for malignant neoplasm of breast
- D05 for carcinoma in situ of breast.

Tumour histology is not used as a variable for risk adjustment and is therefore not analyzed in this section. Instead, however, several authors used the tumour differentiation (also known as grade) for risk adjustment:[221-223]

- Grade I : well differentiated
- Grade II : moderately differentiated
- Grade III : poorly differentiated
- Grade IV : undifferentiated
- Unknown differentiation.

For a detailed description of TNM classification and TNM stages for breast cancer, we refer to KCE report 63.[224, 225]

Table 5.52 shows that the mean age of the 9 633 patients selected in the MCD was 60.1 years. Median age is 60.0 years. Carcinoma in situ represents only 8.5% of all patients operated for breast cancer. Grade II and III are the most represented (34% and 32%, respectively). 41% of the operated patients have grade I breast cancer. Stage 2 disease comes second with 37.8%.

Table 5.52: Breast cancer surgery: Patient and tumour characteristics

		Number	%
Age (years)			
	mean/median	60.1/60	
	Std	12.9	
	Total	9633	
Tumour location			
	Malignant neoplasm of breast	7705	91.55
	Carcinoma in situ	711	8.45
	Total	*8 416*	
Tumour differentiation (grade)			
	Grade I	1255	14.91
	Grade II	2902	34.48
	Grade III	2728	32.41
	Grade IV	88	1.05
	Unknown	1443	17.15
	Total	*8 416*	
Tumour stage			
	0	567	7.42
	I	3 118	40.79
	II	2 885	37.75
	III	960	12.56
	IV	113	1.48
	Total	*7 643*	

Note that these characteristics are based on hospital stays. This implies that a patient who is admitted more than once, will contribute twice (or more) to these characteristics.

5.5.4 Definition of volume

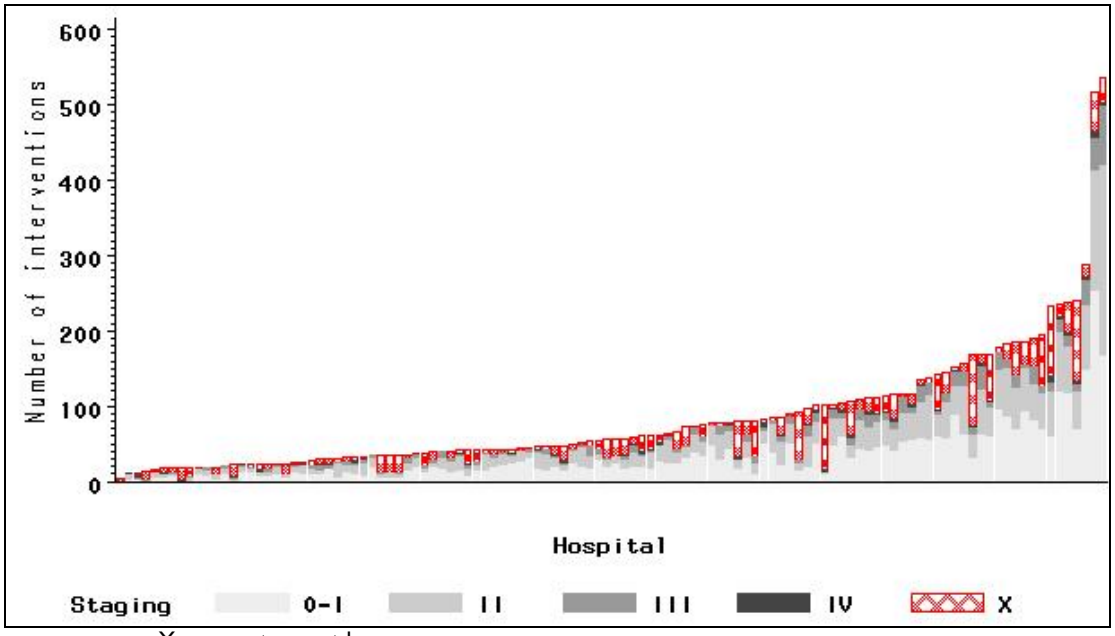
These 9 633 procedures were performed in 114 centres by 805 surgeons. Three quarters of the hospitals performed 109 procedures or less.

Table 5.53: Breast cancer surgery: Summary measures of volume per hospital and per surgeon

	Number	mean	min	25th Pctl	50th Pctl	75th Pctl	Maximum
Hospitals	114	84.5	3.0	33.0	56.0	109.0	537.0
Surgeons*	805	11.6	1.0	2.0	5.0	12.0	316.0

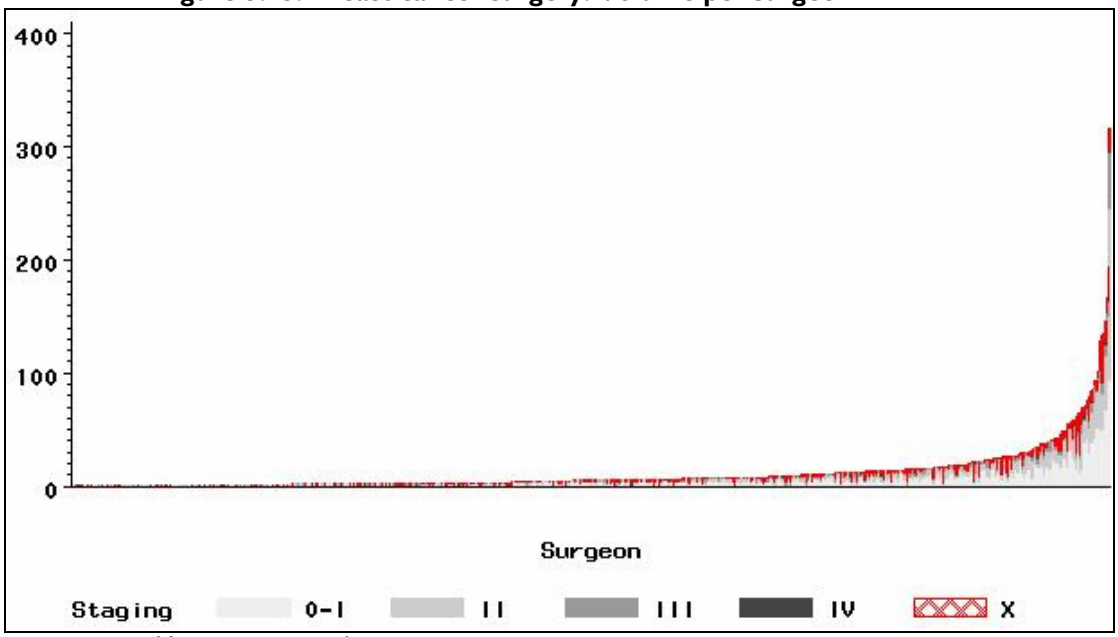
* For 292 procedures the information on surgeon ID was missing.

Figure 5.22: Breast cancer surgery: Volume per hospital



X means stage not known.

Figure 5.23: Breast cancer surgery: Volume per surgeon



X means stage not known.

Guller et al. performed stratified analyses based on whether breast-conserving or breast-ablative therapy was performed.[219] These categories correspond with the following APR-DRGs in the MCD:

- APR-DRG 363 "Breast procedures except mastectomy" = breast-conserving therapy (BCT);
- APR-DRG 362 "Mastectomy procedures" = breast-ablative therapy (BAT).

5.5.5 Definition of outcomes

5.5.5.1 Mortality

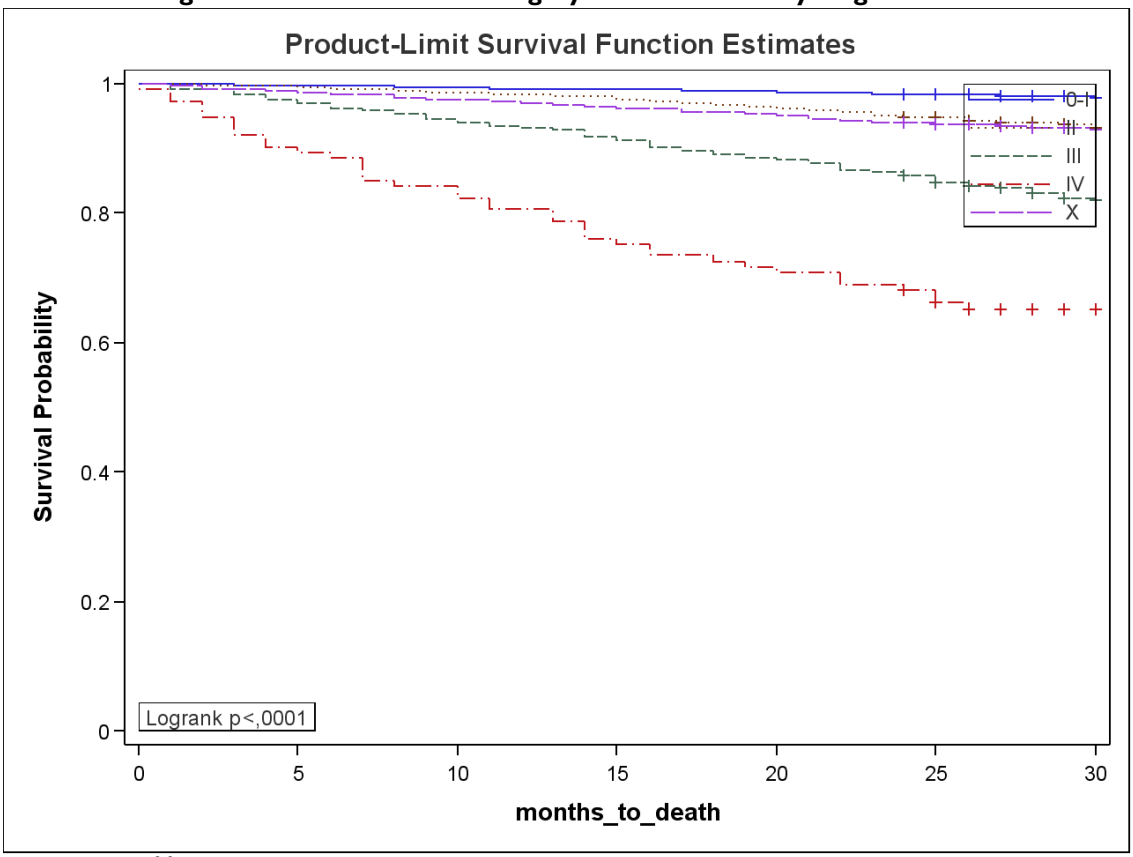
Table 5.54 shows that in-hospital and approximate 30-day mortality is 0.25% after breast cancer surgery (24 deaths). Survival 1 year after surgery is 97.6%. After 2 years it is 94.8%. Mean survival time (median not defined) was 34 months (not visible in Table 5.54).

Table 5.54: Breast cancer surgery: Mortality results

	Number	Number deaths	% deaths	Survival
In-hospital and approximate 30-day mortality	9 450	24	0.25	
1-year mortality	9 450	223	2.4	97.6
2-year mortality	9 450	488	5.2	94.8

Figure 5.23 presents survival after breast cancer surgery, per stage.

Figure 5.24: Breast cancer surgery: Survival curve by stage



X means stage missing.

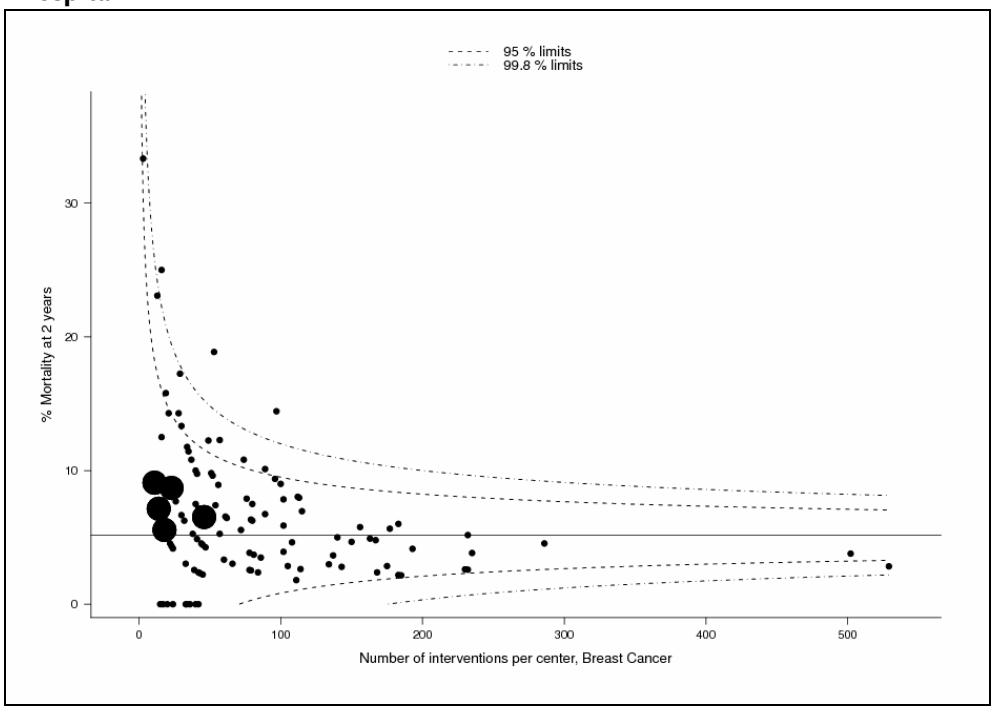
5.5.6 Volume outcome relationship

5.5.6.1 Analysis by hospital

2-YEAR MORTALITY RATE, BY HOSPITAL

Figure 5.24 presents the funnel plot of the number of breast cancer procedures per centre and the 2-year mortality (observed, i.e. without risk adjustment). The size of the point is proportional to the number of hospitals with the same volume and the same outcome. The horizontal line represents the overall 2-year mortality i.e. 5%. There are only 3 centres that are above the 99.8% limits of variability.

Figure 5.25: Breast cancer surgery: Funnel plot of 2-year mortality rate, by hospital



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

Table 5.55 presents the differences in case mix and outcomes for the 114 centres, divided in groups of five based on their annual volume. The international cut off of 10 procedures per year (see discussion section) was not retained as only one centre did not reach this cut off. Volume thresholds were defined by increases of 50 interventions (25 for small volume centres).

Patients treated in low volume centres have a worse case mix:

- patients are older in low volume hospitals (mean age is 62.4 compared to 59.2 in very high volume centres);
- patients have higher levels of co morbid illness in low volume hospitals (4.0% with at least 3 co morbidities versus 2.2% national mean);
- patients in low volume hospitals have a higher likelihood of undergoing breast-ablative surgery (67%) than those in high volume centres (52%);
- patients in low volume hospitals have more stage III and IV patients (together 22.8%) than in the highest volume centres (12.8%).

However, with respect to tumour stage and differentiation, we noticed that low volume centres also have a much higher percentage of missing data than high volume centres:

- 30% (low-volume) versus 16.5% (high volume) missing data for stage;
- 37% (low-volume) versus 22% (high volume) missing data for tumour differentiation.

Finally, 2-year mortality is higher in very small volume centres (8.0%) than in very high volume centres (3.8%).

Table 5.55: Breast cancer surgery: Differences in case-mix and outcomes by hospital volume

		Hospital volume					
		1-25/yr	26-50/yr	51-100/yr	101-150/yr	> 150/yr	All
Number centres		21	32	27	16	18	114
number stays		373	1202	1895	1834	4146	9450
Age	Mean	62.4	61.3	61.0	60.1	59.2	60.1
	Std	14.1	13.5	12.6	12.9	12.6	12.9
Charlson ≥ 3	%	4.0	2.1	2.8	2.1	1.7	2.2
Charlson	Mean	0.4	0.2	0.3	0.3	0.2	0.2
	Std	0.9	0.7	0.8	0.7	0.6	0.7
Principal Diagnosis							
breast neoplasm	%	93.8	93.5	95.8	93.6	93.3	93.9
in situ	%	6.2	6.5	4.2	6.4	6.7	6.1
Type of surgery (APR-DRG)							
breast-ablative therapy	%	66.8	56.7	59.6	54.0	51.7	55.0
breast-conserving therapy	%	33.2	43.3	40.4	46.0	48.3	45.0
LOS	Mean	7.4	6.2	6.6	6.1	5.6	6.0
	Median	7.0	6.0	6.0	5.0	5.0	5.0
	Std	6.0	5.4	5.1	5.3	4.3	4.9
In hospital + 30 day mortality	%	0.8	0.2	0.3	0.3	0.2	0.3
2-year mortality	%	**8.0**	**6.2**	**6.9**	**5.2**	**3.8**	**5.2**
Missing data BCR							
stage	%	29.5	20.4	23.0	18.2	16.5	19.1
tumour differentiation	%	36.5	27.9	31.2	26.6	22.4	26.2
Stage							
0-I	%	40.3	45.2	47.7	49.3	49.4	48.2
II	%	36.9	39.7	38.7	35.5	37.8	37.7
III	%	19.4	13.6	12.1	13.3	11.6	12.6
IV	%	3.4	1.5	1.5	1.8	1.2	1.5
Tumour differentiation (Grade)							
1	%	21.5	19.5	20.3	18.5	16.2	18.0
2	%	50.6	41.9	41.4	40.6	41.4	41.6
3	%	27.0	37.7	37.4	40.3	40.6	39.1
4	%	0.8	0.9	0.9	0.6	1.8	1.3

RELATION BETWEEN HOSPITAL VOLUME AND 2-YEAR MORTALITY RATE

There was an inverse association between the annual volume of the centre and 2-year mortality. The odds of 2-year death decreased by 3.3% 95% CI (-4.35%, -2.27%) when the volume increased by 10%. The magnitude of this association was slightly decreased when patient and tumour characteristics were taken into account, indicating that, although small volume centres tend to have a worse case mix, this is not sufficient to explain entirely the inverse association between volume and outcome.

Table 5.56: Breast cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality

Model without adjustment for case mix				
Factor		Estimate /OR	95%CI	
Hospital volume	increase of 10%	-3.31	-4.35	-2.27
Model with adjustment for patient (MCD data)				
Hospital volume	increase of 10%	-2.55	-3.69	-1.42
Age	increase of 1 year	1.05	1.03	1.06
Principal Diagnosis	Malignant neoplasm of female breast	3.05	1.45	6.38
	Carcinoma in situ of breast	1.00	1.00	1.00
Type of surgery	Breast-ablative surgery	1.83	1.43	2.33
	Breast-conserving therapy	1.00	1.00	1.00
Charlson Score		1.57	1.34	1.83
Model with adjustment for patient and tumour characteristics (MCD and BCR data)				
Hospital volume	increase of 10%	-2.39	-3.64	-1.14
Age	increase of 1 year	1.04	1.03	1.06
Principal Diagnosis	Malignant neoplasm of female breast	1.98	0.95	4.10
	Carcinoma in situ of breast	1.00	1.00	1.00
Type of surgery	Breast-ablative surgery	1.32	1.03	1.71
	Breast-conserving therapy	1.00	1.00	1.00
Charlson score		1.52	1.29	1.79
Stage	unknown	2.93	1.93	4.43
	IV	19.39	11.70	32.13
	III	6.24	4.20	9.25
	II	2.24	1.59	3.16
	0-I	1.00	1.00	1.00
Tumour differentiation (Grade)	unknown	1.72	0.95	3.12
	4	5.35	2.37	12.08
	3	2.95	1.77	4.92
	2	1.06	0.62	1.81
	1	1.00	1.00	1.00

In the previous table, (log of) volume was included in the model as a continuous variable. Sensitivity analyses have been performed using the volume categories from Table 5.55 and show very similar results: the higher mortality from low volumes centres is partially explained by the case mix, but the inverse association remains.

Table 5.57: Breast cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality: sensitivity analysis, centre volume is categorized

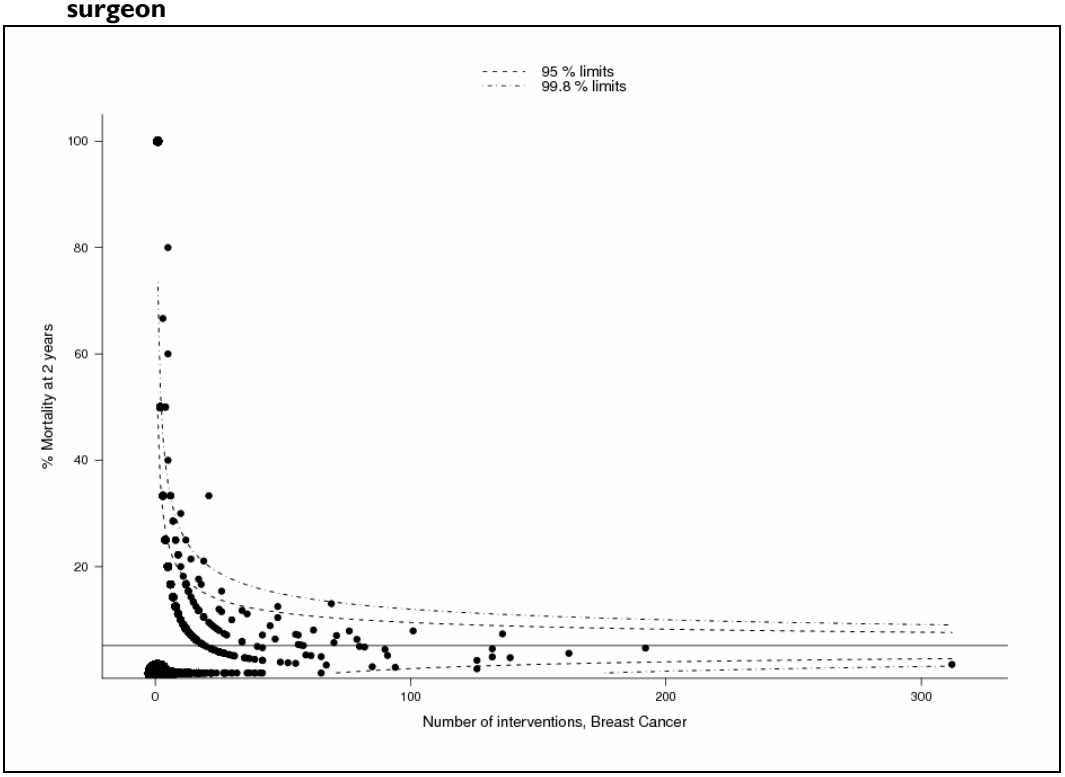
Model without adjustment for case mix				
Factor		OR	95%CI	
Hospital volume	1-25/year	2.19	1.46	3.29
	26-50/year	1.64	1.22	2.21
	51-100/year	1.86	1.42	2.44
	101-150/year	1.37	1.04	1.81
	> 150/year	1.00	1.00	1.00
Model with adjustment for patient (MCD data)				
Hospital volume	1-25/year	1.61	1.03	2.51
	26-50/year	1.43	1.06	1.92
	51-100/year	1.59	1.19	2.13
	101-150/year	1.28	0.93	1.75
	> 150/year	1.00	1.00	1.00
Model with adjustment for patient and tumour characteristics (MCD and BCR data)				
Hospital volume	1-25/year	1.54	1.01	2.35
	26-50/year	1.43	1.02	1.99
	51-100/year	1.62	1.23	2.14
	101-150/year	1.21	0.90	1.63
	> 150/year	1.00	1.00	1.00

5.5.6.2 Analysis by surgeon

2-YEAR MORTALITY RATE, BY HOSPITAL

Figure 5.25 presents the funnel plot of the number of breast cancer procedures per surgeon and the 2-year mortality (observed, i.e. without risk adjustment). The size of the point is proportional to the number of surgeons with the same volume and the same outcome. The horizontal line represents the overall 2-year mortality i.e. 5%. This graphic is not really helpful to assess the volume outcome relationship, as the few low volume centres with extremely high mortality tend to "compress" the graphic.

Figure 5.26: Breast cancer surgery: Funnel plot of 2-year mortality rate, by surgeon



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE SURGEON

Table 5.58 presents the differences in case mix and outcomes per surgeon volume. In contrast with hospital volume, patients treated by low volume surgeons do not have a different case mix. Nevertheless, 2-year mortality is 7.4% for very small volume surgeons and 4.1% for very high volume surgeons.

Table 5.58: Breast cancer surgery: Differences in case-mix and outcomes by surgeon volume

		Surgeon Volume					
		1-2/yr	**3-5/yr**	**6-10/yr**	**11-50/yr**	**>51/yr**	**All**
Number surgeons		267	164	145	192	34	802
Number stays		363	618	1073	3895	3213	9162
Age	Mean	61.1	61.0	60.7	60.6	59.0	60.1
	Std	14.2	13.2	12.9	12.9	12.6	12.9
Charlson ≥ 3	%	2.2	1.8	2.1	2.2	2.2	2.2
Charlson	Mean	0.2	0.2	0.2	0.3	0.2	0.2
	Std	0.7	0.6	0.7	0.7	0.7	0.7
Principal diagnosis							
Breast neoplasm	%	92.0	93.9	94.0	94.5	94.3	94.2
In situ	%	8.0	6.1	6.0	5.5	5.7	5.8
Type of surgery (APR-DRG)							
breast-ablative therapy	%	55.4	54.5	52.6	53.7	59.7	55.8
breast-conserving therapy	%	44.6	45.5	47.4	46.3	40.3	44.2
LOS	Mean	7.4	6.4	6.4	6.1	5.8	6.1
	Median	6.0	6.0	6.0	6.0	5.0	5.0
	Std	8.8	4.0	5.2	4.1	5.2	4.9
In hospital + 30 day mortality	%	1.4	.	0.3	0.2	0.2	0.3
2 year mortality	%	**7.4**	**6.8**	**6.4**	**5.2**	**4.1**	**5.2**
Missing data in BCR							
stage	%	25.9	25.9	24.7	17.5	17.0	19.1
tumor diff	%	36.4	35.6	32.4	24.9	22.5	26.1
Stage							
0-I	%	52.8	48.7	49.3	46.6	48.0	47.7
II	%	30.9	39.3	34.4	38.5	39.2	38.1
III	%	14.1	10.3	13.6	13.7	11.4	12.7
IV	%	2.2	1.7	2.7	1.1	1.4	1.5
Tumour differentiation (Grade)							
1	%	17.3	21.4	18.2	19.6	15.3	17.9
2	%	47.2	43.2	39.9	41.4	41.8	41.7
3	%	34.2	33.7	41.0	37.7	41.8	39.2
4	%	1.3	1.8	1.0	1.3	1.2	1.2

RELATION BETWEEN SURGEON VOLUME AND 2-YEAR MORTALITY RATE

Results from logistic regression taking into account volume of centres and volume of surgeons (with GEE adjustment for clustering of data within hospitals) are shown in Table 5.59. These results show negative estimates of surgeon and centres volume (indicating an inverse relationship between volume and 2-year mortality), and a stronger effect of hospital volume than surgeon volume.

Table 5.59: Breast cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality

Model without adjustment for case mix					
Factor		Effect	95% CI		P value
Volume of surgeon	increase of 10%	-0.87	-1.81	0.07	0.0952
volume of hospital	increase of 10%	-2.64	-3.92	-1.35	0.0005
Model with adjustment for patient (MCD data)					
Volume of surgeon	increase of 10%	-1.03	-2.04	-0.03	0.0663
volume of hospital	increase of 10%	-1.73	-3.07	-0.40	0.0209
Model with adjustment for patient and tumour characteristics (MCD and BCR data)					
Volume of surgeon	increase of 10%	-0.92	-1.98	0.15	0.1180
volume of hospital	increase of 10%	-1.71	-3.09	-0.33	0.0259

5.5.7 Discussion

5.5.7.1 External validation of the definition of the procedure

The Belgian Cancer Registry registered 9 455 new cases of breast cancer in Belgium in 2004. See Table 5.1 (page 45). This number corresponds well to the 9 166 patients with the principal diagnosis of these cancers which were identified in the 2004 Minimal Clinical Data (see Table 5.49, page 100).

5.5.7.2 Summarized results of literature review

The systematic literature search identified 6 systematic reviews in which the volume outcome association (VOA) for breast cancer surgery was studied.[1, 5, 60, 62, 64, 69] These systematic reviews were based on 7 primary studies.[220, 221, 226-230] An additional search for more recent studies which were published in the period 2004-2009 resulted in 7 additional primary studies.[48, 219, 222, 223, 231-233] This brings us to a total of 14 primary studies that analyzed the volume outcome association for breast cancer surgery. For more insight in how these studies were retrieved see the Supplement.

On the basis of the systematic reviews it was concluded in Chapter 2 (see Table 2.2 on page 19) that the limited evidence did not allow a conclusion on the volume-outcome association for breast cancer surgery.

Two minimal hospital volume thresholds were retained from the literature search: a lower threshold of 10 breast cancer surgery procedures per annum and an upper threshold of 150 procedures per annum. These thresholds correspond with the thresholds in the systematic reviews and with the threshold which is currently in use for the recognition of Belgian breast clinics i.e. 150 new diagnoses of breast cancers a year.[234]

Detailed results of the literature review are described in the evidence tables in Appendix 8.

5.5.7.3 *Comparative analysis of literature and Belgian data*

DEFINITION OF VOLUME

In 2004, only 18 out of 114 Belgian hospitals reached the threshold of 150 breast cancer procedures (see Table 5.55). However, this comparison is not entirely correct since the recognition threshold is about new diagnoses of breast cancer while we assessed the number of operations.[234] In addition, it has to be acknowledged that in the transition year 2004 the threshold was defined at only 100 cases per year.

OUTCOME

In this KCE study (see Table 5.54), in-hospital and approximate 30-day mortality is 0.25% after breast cancer surgery. Guller et al. reported an in-hospital mortality of 0.1%, but they excluded patients with radical mastectomy or with metastatic disease.[219] Two-year survival rates were not found in the studies we selected.

Note that some authors calculated the survival period for each patient as the time difference between the date of diagnosis and the date of death or censoring.[221, 222] This contrasts with our method that starts from the date of cancer surgery.

PATIENT CASE MIX

Table 5.60 compares the patient and tumour characteristics of our study population with others. Differences in patient selection can probably explain part of the dissimilarities.

- Remacle et al. used data from the Alliance of Christian Sickness Funds. [235] Their patient selection criteria were very similar to ours: breast cancer diagnosis in the Cancer Registry or billing of breast surgery. When we compare their NIHDI procedure (billing) codes, however, we have to notice that they did not include the billing codes 227021 and 227043 for local exeresis of the tumour (see Table 5.48). This way, they probably selected less stage 0 or I diseases.
- Bailie et al. selected breast cancer patients with and without breast surgery.[222] This could explain why they counted slightly more stage 4 disease since breast cancer surgery is mostly not performed in patients with a very advanced stage of the breast cancer.

Table 5.60: Breast cancer surgery: Patient and tumour characteristics: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Remacle 2007[235]	Bailie 2007[222]
Country	Belgium	Belgium	Northern Ireland
Study period	2004	1998-2003	1996
Age (mean)	60 years	60 years	Not mentioned
Tumour stage			
0-I	48.2%	37%	39.4%
II	37.8%	47%	43.6%
III	12.6%	10%	10.9%
IV	1.5%	6%	6.1%

The following characteristics were additionally used for risk adjustment in other studies but could not be applied in the KCE study because they are not available in MCD or BCR data:

- race;[219, 223]
- median household income;[219, 223]
- socio-economic status;[221, 222]
- college education;[219]
- menopausal status of patient;[219]
- oestrogen receptor and progesterone receptor status;[223]
- tumour size in mm;[222, 223]

- nodal involvement;[219, 221, 223]
- treatments received: radiotherapy, chemotherapy, hormonal therapy;[221, 222]
- location of hospital i.e. rural or urban, teaching or nonteaching;[219]
- time period.[221]

In our study, patients who were treated in low-volume hospitals were more likely to undergo mastectomy: 67% in low-volume centres versus 52% in high-volume centres. Guller et al. mentioned similar statistically significant differences in undergoing BCT between low- and high-volume hospitals. They pointed out that these differences are important for patients since BCT is considered equally efficient regarding survival in the treatment of early-stage breast cancer while it has better results with respect to body image, sexual function and quality of life.[219] With respect to surgeon volume, on the other hand, there were no differences in the type of surgery (see Table 5.58). Stefoski et al. made the same observation in their data.[221]

Although the use of adjuvant treatment could not be analyzed in our study, it is interesting to mention that Stefoski et al. found that chemotherapy rates increased from 12% in very low volume surgeons to 21% in the high category.[221] Remacle et al. found a similar association between adjuvant chemotherapy and hospital volume in the data of the Alliance of Christian Sickness Funds.[235] According to Stefoski et al. the increased use of combined adjuvant therapy by high-volume surgeons strongly suggests a multi-disciplinary approach in the treatment of breast cancer patients.[221]

VOLUME OUTCOME ASSOCIATION

A statistically significant inverse association was found between hospital volume and 2-year mortality after breast cancer surgery. Two-year mortality was 8.0% in very low volume hospitals (<26 procedures/yr) versus 3.8% in high volume hospitals (>150 procedures/yr). Patients operated in low volume hospitals were older, had higher levels of co morbid illness, were more likely to undergo breast-ablative therapy and had a more advanced disease stage. Nevertheless, these differences in the case mix could not explain the survival disadvantage in low volume hospitals.

Bailie et al. reported similar differences in case mix between hospital volume categories.[222] They found that patients treated in higher volume hospitals were younger, had smaller tumours, earlier stage disease, and used more conservative surgery and radiotherapy. Nevertheless, they too found a surgical workload effect on outcome, even after adjustment for clinical and treatment variables.

With respect to the relationship between surgeon volume and survival, 2-year mortality is again higher for very small volume surgeons (7.4% for those with less than 3 procedures per year) than for very high volume surgeons (4.1% for surgeons with more than 51 breast cancer procedures per annum). However, the effect of hospital volume seems stronger than the one of surgeon volume. There were no differences in case mix between the surgeon's volume categories.

We have to note, however, that the extent-of-disease information is missing differentially across volume groups. This implies that inclusion of information on stage and grade in the model could have biased the results.

For high-risk cancer surgery such as oesophagectomy and pancreatectomy, it seems obvious that the attending surgeon has a potential to directly influence mortality; the surgeon's volume represents a component of experience which is likely to affect performance. However, breast cancer surgery is a low-risk procedure where the loco-regional management has limited impact on survival and where patients are unlikely to die from their cancer surgery. Therefore, some authors argue that the outcome of low-risk cancer procedures depends more on the environment and that the relationship with volume is less obvious.[48, 219] Guller et al. hypothesize that high-volume hospitals have a better pre-, peri- and postoperative management of patients.[219] They mention the possibility that the collaboration between physician, nurses, anaesthesiologists, physical therapists and others is better in high-volume centres. This higher level of collaboration within the entire team that cares for the patient probably decreases the in-hospital mortality.

Key points on volume outcome association for breast cancer surgery

- **A total of 10 048 women were hospitalized in 2004 (retrieved in MCD 2004) with a diagnosis of malignant neoplasm of breast cancer. 86% of these women underwent breast-conserving or breast-ablative surgery (mastectomy).**
- **The population studied consisted of those 9 450 stays with breast cancer surgery and whose data could be linked to IMA databases.**
- **In 87% of the cases, information on tumour could be retrieved in the BCR database. Data on stage was available for 79% of stays. In low volume centres (<10/year) this was only 71%.**
- **These interventions were performed in 114 centres by 805 surgeons. Half of the surgeons performed less than 6 procedures per year.**
- **Two-year mortality was 5.2%. Regression models were fitted to assess the association between hospital or surgeon volume on this outcome. The following factors were taken into account in all analyses: age, principal diagnosis (malignant neoplasm or carcinoma in situ, Charlson score (co morbidity), type of procedure (breast-conserving or breast-ablative surgery), tumour stage and tumour differentiation.**
- **The literature review could not provide a clear answer on the volume outcome association for breast cancer surgery; evidence was too limited.**
- **Belgian data showed a statistically significant inverse relationship between volume of centres and 2-year mortality: respectively 8.0% (in very low volume centres, 1-25/year), 6.2% (26-50/year), 6.9% (51-100/year), 5.2% (101-150/year) and 3.8% in very high volume centres (> 150/year). These differences were attenuated but not totally explained by differences in case-mix.**
- **This inverse relationship is also observed with the volume of surgeon: 7.4% 2 year mortality in very low volume surgeons (1-2/year) versus 4.1% in very high volume surgeons (>51/year). However, when accounted for simultaneously, the effect of the centre seems to be more associated with 2-year mortality than the surgeon's volume.**
- **Patients who were treated in low-volume hospitals were more likely to undergo mastectomy: 67% in low-volume centres versus 52% in high-volume centres.**
- **Because data were not retrieved or not available in the databases, the following characteristics could not be used for risk adjustment: use of (neo) adjuvant therapy (chemo- or radiation therapy), acuity of admission (elective versus urgent), intention of surgery (palliative versus curative), menopausal status or oestrogen receptor and progesterone receptor status.**

5.6 LUNG CANCER SURGERY

5.6.1 General description of procedure

As reported by the Belgian Cancer Registry, lung cancer is the second most frequent cancer in males (17%) and the third most frequent in females (7%). It is the most important cause of death by cancer in males (33%).[108] In general, lung cancer surgery is usually limited to early stages of lung cancer that have not spread beyond the lung. It is also performed more often in cases of non-small cell lung cancer, although, in cases of small cell lung cancer, surgery may be performed in conjunction with chemotherapy or radiation.

5.6.2 Primary data selection in MCD 2004

First selection of data was based on ICD-9-CM diagnostic and procedure codes as proposed by Bach, Begg and Sioris, and on NIHDI procedure codes.[129, 236, 237] Note that we deliberately omitted NIHDI procedure 227382 "Thoracotomy with attempted exeresis" because there is no actual lung excision performed during this procedure.

This selection resulted in a total of 29 762 hospital stays but is too broad since it also includes patients with a secondary diagnosis of malignant neoplasm. This implies that patients with lung cancer as pre-existing condition are also included (see section 4.2.1, page 39). The definition of the population studied will be defined in the next section.

Table 5.61: Lung cancer surgery: Primary data selection in MCD 2004

SELECTION 1	SELECTION 2	SELECTION 3
ICD-9-CM (principal or secondary) diagnosis code	**ICD-9-CM procedure code**	**NIHDI procedure code**
1622 Malignant neoplasm of main bronchus	321 Other excision of bronchus	227220 Total or partial lung excision with lymph node dissection for cancer
1623 Malignant neoplasm of upper lobe, bronchus or lung	3229 Other local excision or destruction of lesion or tissue of lung	227242 Radical pneumonectomy with mediastinal lymph node dissection and with ligation of the lung vessels within the pericardial (heart) sac
1624 Malignant neoplasm of middle lobe, bronchus or lung	323 Segmental resection of lung	227264 Total or partial lung excision
1625 Malignant neoplasm of lower lobe, bronchus or lung	324 Lobectomy of lung	227286 Resection of bronchial trunk or trachea via thoracotomy with tracheobronchial or bronchobronchial anastomosis
1628 Malignant neoplasm of other parts of bronchus or lung	325 Pneumonectomy	
1629 Malignant neoplasm of bronchus and lung, unspecified	326 Radical dissection of thoracic structures	
STAYS SELECTED = 28 989 stays	**STAYS SELECTED = 1 913 stays**	**STAYS SELECTED = 1 958 stays**
TOTAL STAYS SELECTED = 29 762 stays (selection 1 OR selection 2 OR selection 3)		

5.6.3 Definition of procedure

5.6.3.1 Primary hospital stays in Minimal Clinical Data

Only 10 701 of the 29 762 selected hospital stays had a principal diagnosis of malignant neoplasm of bronchus and lung. As shown in Table 5.62, 1 313 stays (12.3% of 10 701 stays) corresponding to 1 296 patients (17.6% of 7 360 patients) had a procedure for lung cancer surgery as described above.

Table 5.62: Lung cancer surgery: Principal diagnosis and percentage of cancer surgery (per stay and per patient)

Principal diagnosis	**Total number hospital stays**	**Stays with lung cancer surgery**		**Total number patients**	**Patients with lung cancer surgery**	
		Number	%		Number	%
162 Malignant neoplasm of bronchus and lung	10 701	1 313	12.3	7 360	1 296	17.6

The APR-DRGs of these 1 313 stays are described in Table 5.63.

To create a study population of patients which is as homogenous as possible, we retained only the 1 206 stays (corresponding to 1 192 patients) that are included in the APR-DRG 120 "Major respiratory procedures" and were retained in our analysis.

Table 5.63: Lung cancer surgery: All APR-DRGs of stays with a principal diagnosis of a malignant neoplasm of bronchus and lung AND with surgical resection of the lung

APR-DRG	Number	Percent
004-TRACHEOSTOMY EXCEPT FOR FACE, MOUTH & NECK DIAGNOSES / p3 - P	20	1.52
120-MAJOR RESPIRATORY PROCEDURES / 4 - P	**1206**	**91.85**
121-NON-MAJOR RESPIRATORY PROCEDURES / 4 - P	87	6.63
TOTAL	1313	100.00

Results of definition: 1 206 stays with surgery for resectable lung cancer were selected in the Minimal Clinical Data. These stays concerned 1 192 individual patients and 97 centres.

5.6.3.2 Linkage with data from IMA and BCR

The MCD data were linked with the Common Sickness Funds Agency (IMA) database and the Belgian Cancer Registry (BCR) database. Table 5.64 shows that 98.2% of 1 206 MCD stays could be linked with IMA data and 87.4% with BCR data on the basis of ICD-10 code C34 "Malignant neoplasm of bronchus and lung". Information on tumour stage could be retrieved for 74.8% of 1 206 MCD stays.

Table 5.64: Lung cancer surgery: Percentage of linkage of MCD data with data from IMA and BCR

	Number	**%**
Number of stays in MCD selection	1 206	
Linkage with IMA	1 184	98.2
Linkage with BCR	1 054	87.4
Linkage with BCR and data about stage	902	74.8

Results of linkage: 1 054 patients with lung cancer surgery could be linked with Belgian Cancer Registry data. For 902 of these patients, the BCR contained information on the stage of the tumour.

5.6.3.3 Patient and tumour characteristics

The two main histological types are small cell lung cancer and non-small cell lung cancer (NSCLCa).[108] In accordance with the histological groups defined by the International Agency for Research on Cancer (IARC) these two types are defined by the following ICD-O-3 codes:[125]

- small cell lung carcinoma: 8041-8045, 8246
- non-small cell lung cancer (NSCLCa) comprises the following groups:
  - squamous cell carcinoma: 8050-8078, 8083-8084
  - adenocarcinoma: 8140, 8211, 8230-8231, 8250-8260, 8323, 8480-8490, 8550-8551, 8570-8574, 8576
  - large cell carcinoma (include giant cell, clear cell and large cell undifferentiated carcinoma): 8010-8012, 8014-8031, 8035 and 8310.

A description of TNM classification and TNM stages for lung cancer were described by Lababede et al.[238]

As shown in Table 5.65, mean age of patients was 64.5 years (median is 66 years) and 77% of them were men. The predominant histological type of tumour was non-small cell lung (88.7%). Out of 902 operated patients with available stage information, more than half (56%) had stage 0 or I lung cancer.

Table 5.65: Lung cancer surgery: Patient and tumour characteristics

		Number	%
Age (years)			
	Mean/median	64.5/66.0	
	Std	10.1	
	Total	1206	
Sex			
	Male	931	77.2
	Female	275	22.8
	Total	1206	
Tumour histology			
	Small cell lung carcinoma	23	2.2
	Non-small cell lung cancer	934	88.6
	Other	97	9.2
	Total	1 054	
Tumour stage			
	I	505	56.0
	II	185	20.5
	III	185	20.5
	IV	27	3.0
	Total	*902*	

5.6.4 Definition of volume

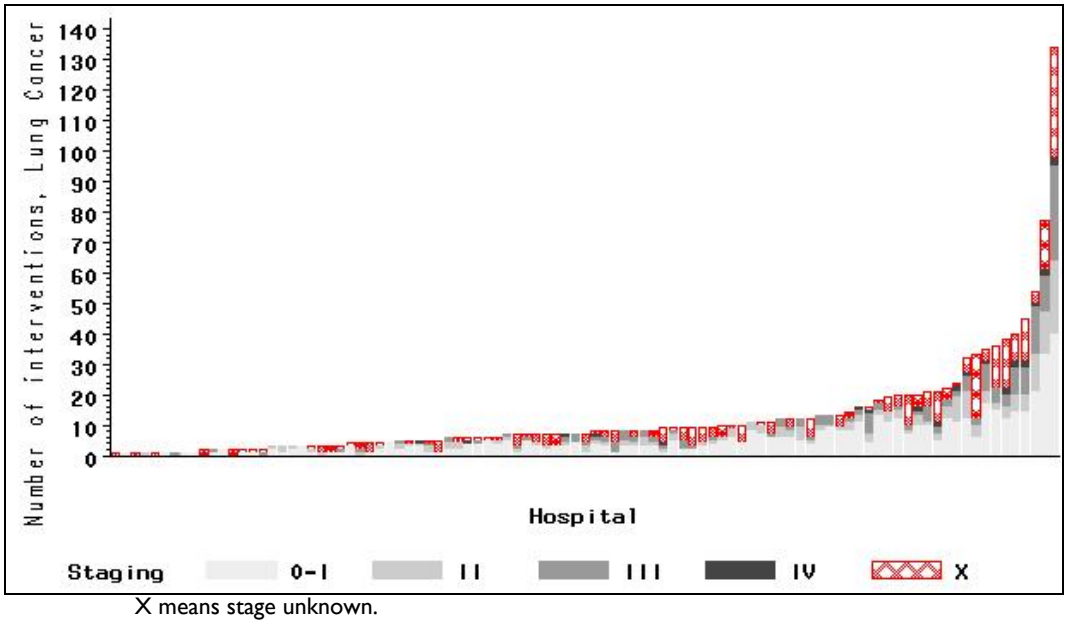
These 1 206 procedures were performed in 97 centres by 154 surgeons. Three quarters of the hospitals performed 13 procedures or less. Three quarters of the surgeons performed 8 procedures or less. 18 surgeons (12% of all surgeons) operated at more than one hospital.

Table 5.66: Lung cancer surgery: Summary measures of volume per hospital and per surgeon

	Number	mean	min	25th Pctl	50th Pctl	75th Pctl	Maximum
Hospitals	97	12.4	1.0	4.0	7.0	13.0	134.0
Surgeons*	154	7.7	1.0	2.0	4.0	8.0	106.0

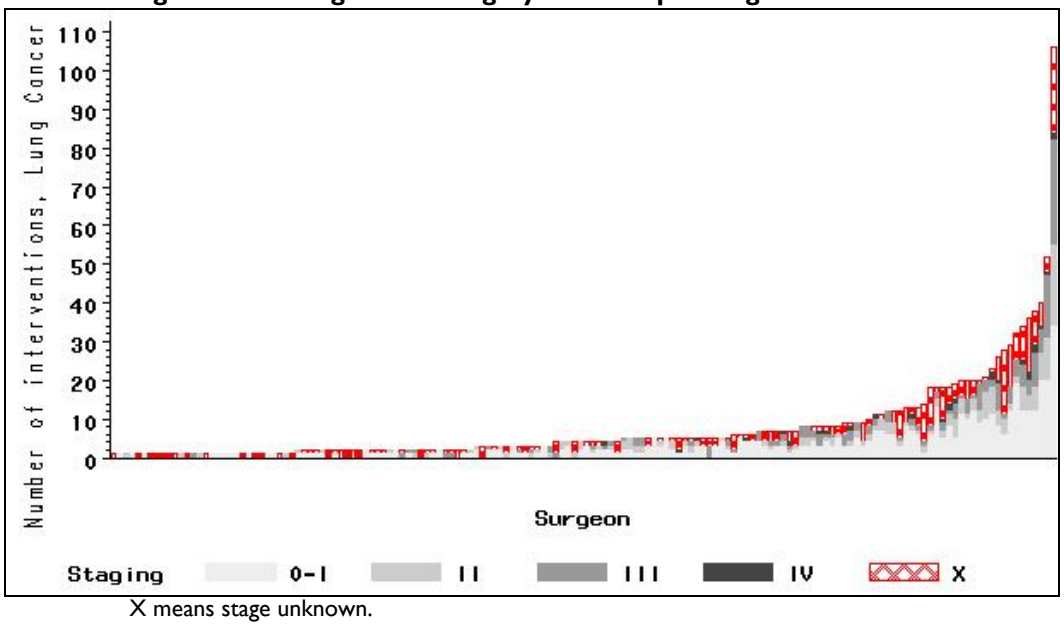
* For 21 procedures (1.7%) information on surgeon was missing.

Figure 5.27: Lung cancer surgery: Volume per hospital



X means stage unknown.

Figure 5.28: Lung cancer surgery: Volume per surgeon



X means stage unknown.

5.6.5 Definition of outcomes

5.6.5.1 Mortality

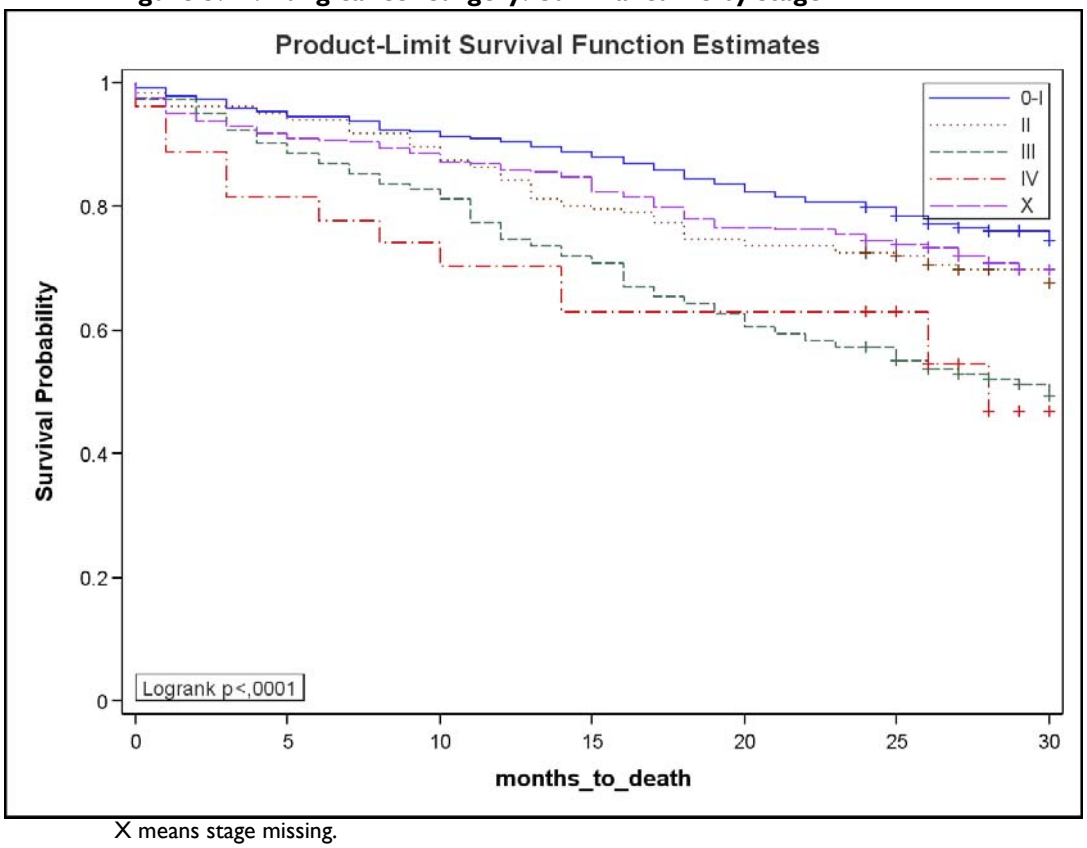
Table 5.67 shows that in-hospital and approximate 30-day mortality is 3.8% after lung cancer surgery. Survival 1 year after surgery is 85.5%. After 2 years it is 73.5%. Mean (median not defined) survival time was 27 months (not shown in Table 5.67).

Table 5.67: Lung cancer surgery: Mortality results

	Number	Number deaths	% deaths	Survival
In-hospital and approximate 30-day mortality	1184	45	3.80	
1-year mortality	1184	172	14.5	85.5
2-year mortality	1184	314	26.5	73.5

Figure 5.28 confirms how survival improves with decreasing disease stage. Patients with stage 3 and 4 lung cancer have the worst survival curves while stage 1 and 2 patients have the best survival. Patients with unknown disease stage, i.e. the purple line in Figure 5.28, are in between the first two stage groups.

Figure 5.29: Lung cancer surgery: Survival curve by stage



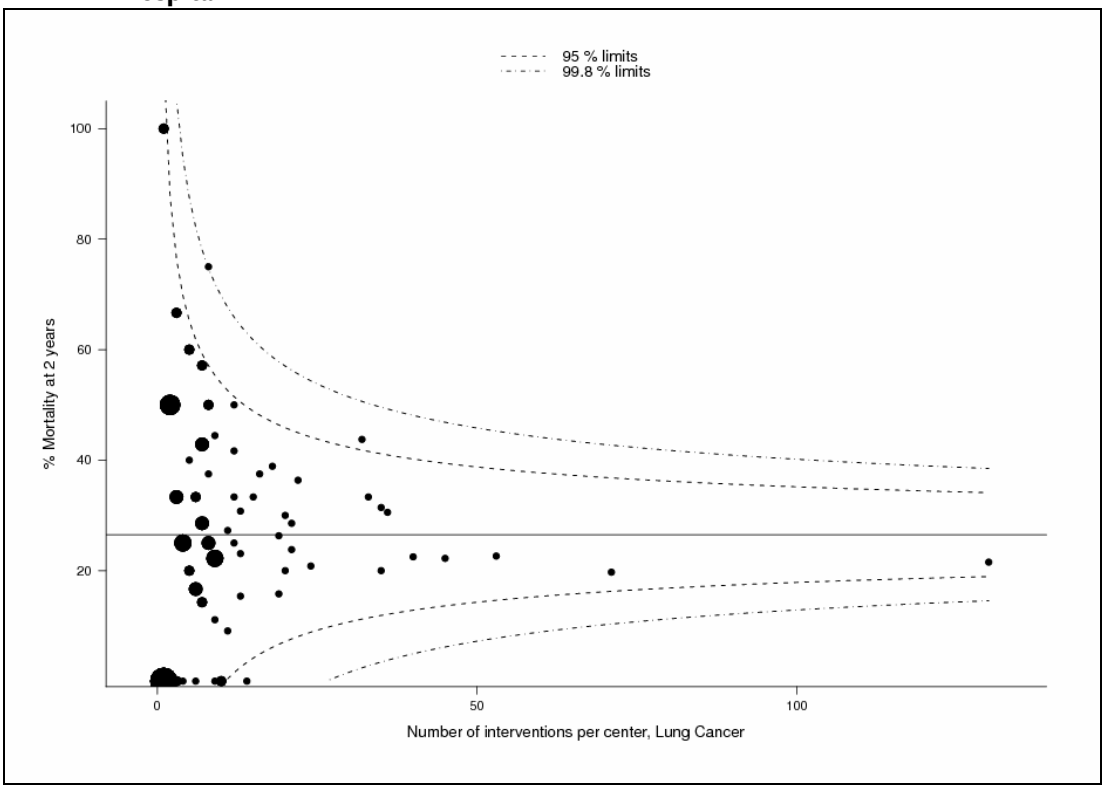
X means stage missing.

5.6.6 Volume outcome relationship

5.6.6.1 Analysis by hospital

Figure 5.29 presents the funnel plot of the number of lung cancer surgeries and the 2-year mortality (observed, i.e. without risk adjustment). The size of the point is proportional to the number of hospitals with the same volume and the same outcome. The horizontal line represents the overall 2-year mortality i.e. 26.5%. None of the hospitals are outliers above the 99.8% limits of variability.

Figure 5.30: Lung cancer surgery: Funnel plot of 2-year mortality rate, by hospital



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

Table 5.68 presents differences in case mix and outcome based on hospital volume. The smallest volume category is based on the 25th percentile (4 interventions/year). The four other volume categories are chosen arbitrary by increases in volume of 5, 10 and 20 interventions. Age and sex are not associated with volume. More patients with Charlson score ≥ 3 and with APR-DRG severity 3 or 4 are present in very small volume centres (<4/year). On the other hand, there are more patients with stage 0-I in very small centres (68%) than in all hospitals (56%) This should be counter balanced by the fact that more data on stage are missing in very small centres (33% versus 24% in all hospitals). Histology seems also a bit different i.e. the category 'other' represents 15% in small volume hospitals versus 9% in all centres). Two-year mortality is around 26 to 29% in all hospitals except in the 4 high volume centres (>40/year), where it drops to 21.4%.

Table 5.68: Lung cancer surgery: Differences in case-mix and outcomes by hospital volume

		Hospital volume					
		1-4/yr	5-10/yr	11-20/yr	21-40/yr	> 40/yr	All
Number centres		29	26	18	10	4	97
Number stays		66	261	259	299	299	1184
Age	Mean	64.1	65.6	64.8	63.8	64.1	64.5
	Std	11.7	9.7	10.0	10.1	10.3	10.1
Gender (Male)	%	78.8	77.0	80.7	73.2	77.9	77.2
Charlson ≥ 3	%	15.2	10.0	11.2	9.0	12.4	10.9
Charlson	Mean	1.2	1.0	1.0	1.0	0.9	1.0
	Std	1.3	1.1	1.3	1.1	1.3	1.2
Severity index (APR-DRG)							
1	%	15.2	10.7	14.7	14.4	12.7	13.3
2	%	33.3	44.1	41.7	52.2	44.8	45.2
3	%	34.8	31.4	25.5	24.4	30.8	28.4
4	%	16.7	13.8	18.1	9.0	11.7	13.2
LOS	Mean	13.8	16.6	17.2	14.0	14.5	15.4
	Median	12.5	12.0	14.0	11.0	12.0	12.0
	Std	6.1	14.6	10.9	8.4	8.7	10.6
In hospital + 30 day mortality	%	6.1	3.1	3.9	4.7	3.0	3.8
2 year mortality	%	**27.3**	**29.5**	**26.3**	**29.1**	**21.4**	**26.5**
Missing stage	%	33.3	28.4	15.8	28.8	19.7	23.8
Missing histology	%	16.7	14.2	8.1	12.0	8.4	11.0
Stage							
0-I	%	68.2	58.8	65.6	53.5	45.0	56.0
II	%	20.5	19.3	18.3	20.7	23.3	20.5
III	%	11.4	19.3	14.2	21.1	28.3	20.5
IV	%	0	2.7	1.8	4.7	3.3	3.0
Histology							
non small cell lung	%	83.6	89.7	88.2	88.2	89.4	88.6
small cell lung	%	1.8	1.8	0.8	3.4	2.6	2.2
other	%	14.5	8.5	10.9	8.4	8.0	9.2

RELATION BETWEEN HOSPITAL VOLUME AND 2-YEAR MORTALITY RATE

Results from logistic regressions (with GEE adjustment for correlations of patients within centres) are shown in Table 5.69. There is a statistically significant inverse relationship between the (log of) volume of centres and 2-year mortality. This association is stronger when the stage is taken into account (as small centres have proportionally more patients with stage I patients).

Table 5.69: Lung cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality

Model without adjustment for case mix					
		Effect	95%CI		
Hospital Volume	increase of 10%	-1.27	-2.05	-0.49	
Model with adjustment for patient (MCD data)					
Hospital Volume	increase of 10%	-1.04	-1.87	-0.21	
		Odds Ratio	95%CI		
Age	increase of 1 y.	1.02	1.00	1.03	
Sex	male vs female	1.29	0.87	1.93	
Charlson score	Increase of 1 point	1.35	1.16	1.58	
Model with adjustment for patient and tumour characteristics (MCD and BCR data)					
Hospital Volume	increase of 10%	-1.74	-2.58	-0.91	
		Odds Ratio	95%CI		
Age	increase of 1 y.	1.02	1.01	1.04	
Sex	male vs female	1.24	0.84	1.82	
Charlson score	Increase of 1 point	1.40	1.17	1.66	
Stage	unknown	1.99	1.36	2.91	
	IV	2.78	1.06	7.27	
	III	3.50	2.53	4.85	
	II	1.64	1.16	2.32	
	0-I	1.00	1.00	1.00	
Histology	unknown	0.44	0.13	1.57	
	non small cell lung	1.09	0.38	3.09	
	other	0.86	0.27	2.71	
	small cell lung	1.00	1.00	1.00	

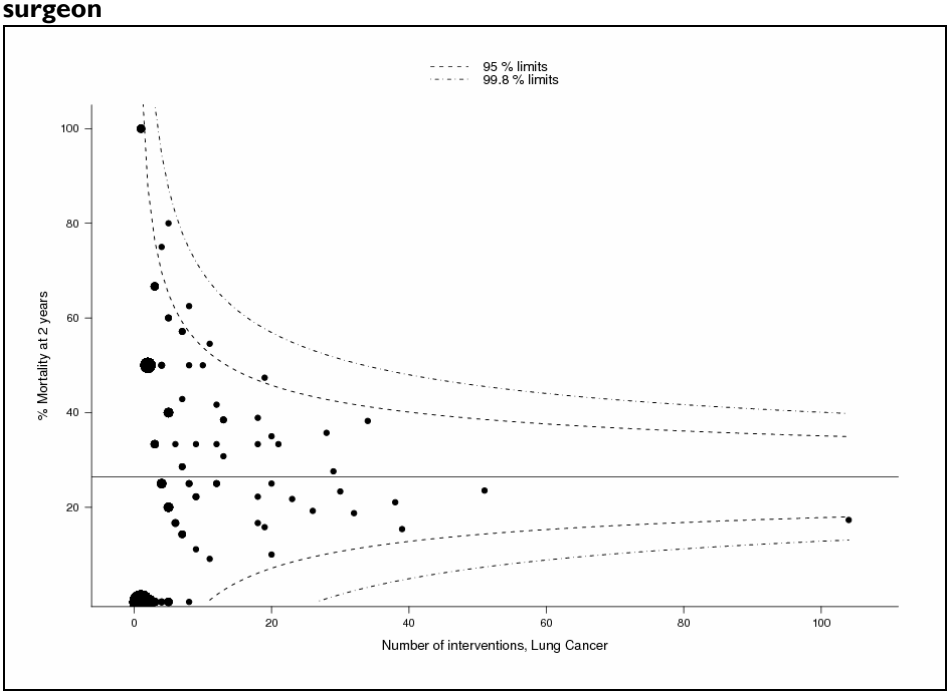
These results were confirmed when the highest volume centre was excluded from the analysis (sensitivity analysis).

5.6.6.2 *Analysis by surgeon*

MORTALITY RATE, BY SURGEON

Figure 5.31 presents the funnel plot of the number of lung cancer surgeries per surgeon and the 2-year mortality (observed, i.e. without risk adjustment). The size of the point is proportional to the number of surgeons with the same volume and the same outcome. The horizontal line represents the overall 2-year mortality i.e. 26.5%. None of the surgeons are outliers above the 99.8% limits of variability.

Figure 5.31: Lung cancer surgery: Funnel plot of 2-year mortality rate, by surgeon



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE SURGEON

Table 5.70 presents differences in case mix and outcome based on volume of surgeon. The smallest volume category is based on the 25th percentile (2 interventions per year). The other categories are defined by increases in volume of 3, 5, 10 and 20 interventions. Age, sex, Charlson score are not associated with volume. There are, however, more patients with stage I in very small volume surgeons (66%) than in all hospitals (56%). This should be counter balanced by the fact that more data on stage are missing in very small volume surgeons (44% versus 24% for all surgeons). Two-year mortality is 21.4% for very small volume surgeons and 23.1% for very high volume surgeons (> 20/year).

Table 5.70: Lung cancer surgery: Differences in case-mix and outcomes by hospital volume

		Surgeon Volume					
		1-2/yr	3-5/yr	6-10/yr	11-20/yr	> 20/yr	All
Number surgeons		48	35	28	20	19	152
Number stays		84	172	171	279	455	1161
Age	Mean	64.0	64.2	64.4	64.7	64.6	64.5
	Std	9.7	10.5	10.3	10.5	9.7	10.1
Male gender	%	75.0	78.5	76.0	73.8	78.5	76.7
Charlson ≥ 3	%	11.9	8.7	10.5	10.4	11.6	10.8
Charlson	Mean	1.0	1.0	1.0	0.9	1.0	1.0
	Std	1.1	1.0	1.1	1.3	1.3	1.2
Severity index (APR-DRG)							
1	%	14.3	13.4	12.9	16.8	11.0	13.3
2	%	45.2	47.1	43.9	44.8	45.5	45.3
3	%	27.4	29.1	29.2	21.9	31.9	28.3
4	%	13.1	10.5	14.0	16.5	11.6	13.1
LOS	Mean	14.7	16.8	15.5	16.3	14.5	15.4
	Median	12.0	12.0	12.0	13.0	12.0	12.0
	Std	9.2	15.1	10.4	11.1	8.4	10.6
In hospital + 30 day mortality	%	2.4	3.5	4.1	5.4	3.1	3.8
2 year mortality	%	**21.4**	**30.2**	**29.2**	**29.4**	**23.1**	**26.4**
Missing data in BCR							
stage	%	44.0	22.1	22.2	25.1	21.3	24.1
histology	%	29.8	6.4	13.5	9.7	9.2	11.0
Stage							
I	%	66.0	59.0	56.4	62.2	49.2	55.7
II	%	21.3	24.6	15.0	19.1	22.6	20.9
III	%	12.8	14.2	24.1	15.8	24.9	20.3
IV	%	.	2.2	4.5	2.9	3.4	3.1
Histology							
non small cell lung	%	89.8	85.1	89.9	88.5	88.9	88.4
other	%	8.5	11.2	10.1	10.7	7.7	9.4
small cell lung	%	1.7	3.7	.	0.8	3.4	2.2

RELATION BETWEEN SURGEON VOLUME AND 2-YEAR MORTALITY RATE

There was a statistically significant inverse relationship between the (log of) volume of surgeon and 2-year mortality, after adjustment for patients and tumour characteristics. However, this association was not robust when the largest outlying surgeon (> 100/year) was removed from the analysis.

Table 5.71: Lung cancer surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of 2-year mortality

			Effect	95%CI	
Volume of surgeon	Increase of 10%	Adjustment for patients and tumour characteristics	-1.55	-2.80	-0.31
Volume of surgeon	Increase of 10%	Sensitivity analysis: after exclusion of 1 high volume surgeon (>100/year)	-0.74	-2.08	0.59

5.6.7 Discussion

5.6.7.1 External validation of the definition of the procedure

The Belgian Cancer Registry registered 6 994 new cases of lung and bronchus cancer in Belgium in 2004. See Table 5.1 (page 45). In the 2004 Minimal Clinical Data (MCD), we identified 7 360 patients with the principal diagnosis of malignant neoplasm of lung or bronchus (see Table 5.62). The higher number of patients in the MCD is explainable by the fact that these hospital stays are not restricted to patients whose cancer started in 2004.

5.6.7.2 Summarized results of literature review

The systematic literature search identified 7 systematic reviews in which the volume outcome association (VOA) for lung cancer surgery was studied.[1, 5, 59, 60, 62, 64, 69] These systematic reviews were based on 11 primary studies.[115, 116, 129, 133, 141, 142, 189, 197, 236, 239, 240] An additional search for more recent studies which were published in the period 2004-2009 resulted in 8 additional primary studies.[94, 145, 153, 237, 241-244] This brings us to a total of 19 primary studies that analyzed the volume outcome association for lung cancer surgery. For more insight in how these studies were retrieved see the Supplement.

On the basis of the systematic reviews it was concluded in Chapter 2 (see Table 2.2 on page 19) that there is evidence for an inverse relation between hospital volume and mortality for lung cancer surgery. This means that mortality decreases when the number of procedures performed by a hospital increases. Such relationship could not be proved between surgeon volume and mortality.

However, the effect of volume on outcome for this procedure is small when compared to pancreatic or oesophageal cancer surgery. Killeen et al. calculated that the number of patients that a high-volume unit would need to treat to prevent one death associated with low volume is between 20 and 50.[69]

In relation to a cut off point to define low and high volume, studies differed widely. Therefore, it was impossible to define one minimal hospital volume threshold.

Detailed results of the literature review are described in the evidence tables in Appendix 8.

5.6.7.3 *Comparative analysis of literature and Belgian data*

DEFINITION OF VOLUME

As mentioned before, it was impossible to define a cut off point for hospitals or surgeon volume on the basis of the literature review.

OUTCOME

Table 5.72 compares Belgian outcome data with those published in the studies that were selected. Although patient selection differs among these studies, short-term mortality rates seem comparable (except from the study by Hannan).

Table 5.72: Lung cancer surgery: Mortality: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Bach 2001 [236]	Hannan 2002 [189]	Urbach 2005 [94]	Sioris 2008 [237]
Country	Belgium	USA	USA	Canada	Finland
Study period	2004	1985-1996	1994-1997	1994-1999	1988-2002
Patient selection	Lung cancer + lung surgery (all types)	NSCLCa stage I to IIIa + lung surgery (all types)	Lung cancer + lung lobectomy	Lung cancer + lung lobectomy or pneumo-nectomy	NSCLCa + lung surgery (all types)
In-hospital or 30-day mortality	3.8%	4%	1.86%	4.2%	5.4%
1-year survival	85.5%				
2-year survival	73.5%	64%			

PATIENT CASE MIX

As mentioned before, Bach et al. included only patients 65 years or older who received a diagnosis of stage I, II or IIIa non-small cell lung cancer.[236] This specific patient population probably explains the differences with our population as shown in Table 5.73.

Table 5.73: Lung cancer surgery: Patient and tumour characteristics: comparison of Volume-Outcome KCE study with other studies

	KCE 2004	Bach 2001 [236]
Country	Belgium	USA
Study period	2004	1985-1996
Age	64.5 yr mean 66 yr median	Not mentioned
Male	77%	59%
Tumour histology		
Small cell lung carcinoma	2.2%	0%
Non-small cell lung cancer	88.6%	100%
Other	9.2%	0%
Tumour stage		
I	56.0%	69%
II	20.5%	20%
III	20.5%	11%
IV	3.0%	0%

In many studies the type of operative procedure is included in the risk adjustment.[115, 133, 236, 237, 243, 244] A distinction is made between the following types of lung cancer surgery:

- wedge resection: a small section of the lung is removed; the tumour is removed along with a small amount of healthy tissue surrounding it;
- lobectomy: surgical removal of a whole part of the lung;
- pneumonectomy: surgical removal of the entire lung.

Sioris et al. illustrated that hospital mortality is 3.4% for lobectomies versus 10.6% for pneumonectomies.[237] Including the type of surgery in our analyses was, however, not possible since the NIHDI procedure codes are not detailed enough (see Table 5.61). In addition, the ICD-9-CM procedure codes in the MCD are not reliable enough to make such a distinction.

The following characteristics were additionally used for risk adjustment in other studies but could not be applied in the KCE study because they are not available in MCD or BCR data:

- race;[115, 133, 153, 236]
- median household income;[115, 133, 153, 236]
- college education;[153]
- acuity of index admission i.e. elective, urgent or emergent;[115, 133, 153]
- result of surgery i.e. curative or incomplete;[237]
- range of lymph node dissection;[244]
- delay to surgery;[237]
- year of surgery;[133, 237]
- use of adjuvant therapy i.e. chemotherapy or radiation therapy.[153, 237]

Key points on volume outcome association for lung cancer surgery

- **A total of 7 360 patients were hospitalized in 2004 (retrieved in MCD 2004) with a diagnosis of malignant neoplasm of bronchus and lung. 18% of these patients underwent surgery.**
- **The population studied consisted of those 1 206 stays with lung cancer surgery and whose data could be linked to IMA databases.**
- **In 87% of the cases, information on tumour could be retrieved in the BCR database. Data on stage was available for 75% of stays. In low volume centres (<1-4/year) this was only 66%. In low volume surgeons (1-2/year); this proportion reached 44%**
- **These interventions were performed in 97 centres by 154 surgeons. In half of the centres less than 8 procedures per year were performed. Half of the surgeons performed less than 5 procedures per year.**
- **Two-year mortality was 27%. Regression models were fitted to assess the association between hospital or surgeon volume on this outcome. The following factors were taken into account in all analyses: age, sex, Charlson score (co morbidity), tumour stage and tumour histology. Type of surgery i.e. lobectomy versus pneumonectomy could not be assessed in our data because NIHDI procedure codes are not specific enough.**
- **Systematic reviews showed evidence for an inverse relationship between hospital volume and mortality for lung cancer surgery i.e. mortality decreases when the number of procedures performed by a hospital increases. Such relationship could not be proved between surgeon volume and mortality.**
- **Systematic reviews did not allow defining one minimal hospital volume threshold because studies differed too widely.**
- **Belgian data showed a statistically significant inverse relationship between hospital volume and 2-year mortality: respectively 27.3% (in very low volume centres, 1-4/year), 29.5% (5-10/year), 26.3% (11-20/year), 29.1% (21-40/year) and 21.4% in four very high volume hospitals (> 40/year). These differences were reinforced when case mix was taken into account, as low volume hospitals treated more patients with stage I then high volume centres.**
- **This inverse relationship was also observed with the volume of surgeon, but the association was mainly driven by one surgeon with a very high volume and a good outcome (>100/year).**
- **Because data were not retrieved or not available in the databases, the following characteristics could not be used for risk adjustment: use of (neo) adjuvant therapy (chemo- or radiation therapy), acuity of admission (elective versus urgent), intention of surgery (palliative versus curative), type of surgical resection (total or partial resection, i.e. lobectomy versus pneumonectomy).**

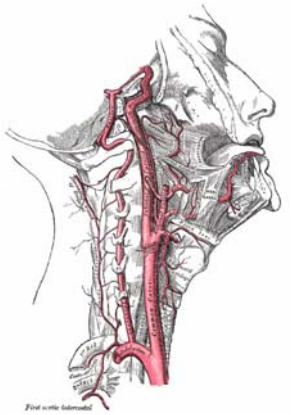
6 RESULTS FOR FOUR CARDIOVASCULAR PROCEDURES

6.1 CAROTID ENDARTERECTOMY (CEA) + CAROTID STENTING (CAS)

6.1.1 General description of procedure

Carotid Endarterectomy (CEA) and Carotid stenting (CAS) are procedures which are undertaken to reduce the risk of stroke in both symptomatic and asymptomatic patients with stenosis (i.e. narrowing) of the carotid arteries in the neck (see Figure 6.1). CEA is the surgical removal of an atheromatous plaque (i.e. a gradual build up of fatty materials) from the inside of the carotid artery to restore blood flow. CAS, on the other hand, involves a metal mesh tube (i.e. stent) which is percutaneously placed in the carotid artery to increase blood flow.[245]

Figure 6.1: The right carotid artery



Source: Gray's Anatomy – Public domain

6.1.2 Primary data selection in MCD 2004

1. CEA was selected by means of the ICD-9-CM procedure codes 3812, 3832, and 3842 as proposed by Christian et al, and/or by the NIHDI code 235071/082 as proposed in previous KCE studies.[93, 246, 247]
2. CAS was selected by means of the ICD-9-CM procedure code 3990 because of lack of a NIHDI billing code.[246, 247]

These selection criteria resulted in 10 890 stays (Table 2.1).

Table 6.1: CEA/CAS: Primary data selection in MCD 2004

SELECTION 1		SELECTION 2	
ICD-9-CM procedure code		NIHDI procedure code	
Code	**label**	**Code**	**label**
3812	Endarterectomy of other vessels of head and neck.	235082	Revascularisation of carotid or vertebral arteries by means of endarterectomy, endoaneurysmorrhaphy, bypass or resection with grafts or anastomosis.
3832	Resection of vessel with anastomosis of other vessels of head and neck.		
3842	Resection of vessel with replacement of other vessels of head and neck.		
3990	Insertion of non-drug-eluting peripheral vessel stent(s).		
STAYS SELECTED = 10 693		**STAYS SELECTED = 3 922**	
TOTAL STAYS SELECTED= 10 890 (selection 1 OR selection 2)			

6.1.3 Definition of procedure

Only 4 030 stays (37%) of these 10 890 selected stays, had an ICD-9-CM principal diagnosis code 433 'Occlusion and stenosis of precerebral arteries'. This proved that the primary selection procedure was not specific enough and had to be refined.

Within ICD-code 433, there appeared to be a need for specification with respect to the carotid artery (Table 6.2).

Table 6.2: CEA/CAS: List of all principal diagnostic ICD-9 codes within ICD-9-CM procedure code 433

	Number (Number=4030)	%
43310 Occlusion and stenosis of carotid artery without cerebral infarction	**3075**	**76.30**
43330 Occlusion and stenosis of multiple and bilateral precerebral arteries without cerebral infarction	535	13.28
43311 Occlusion and stenosis of carotid artery with cerebral infarction	**297**	**7.37**
43331 Occlusion and stenosis of multiple and bilateral precerebral arteries with cerebral infarction	67	1.66
43380 Occlusion and stenosis of other specified precerebral artery without cerebral infarction	32	0.79
43320 Occlusion and stenosis of vertebral artery without cerebral infarction	15	0.37
43300 Occlusion and stenosis of basilar artery without cerebral infarction	5	0.12
43321 Occlusion and stenosis of vertebral artery with cerebral infarction	2	0.05
43381 Occlusion and stenosis of other specified precerebral artery with cerebral infarction	1	0.02
43390 Occlusion and stenosis of unspecified precerebral artery without cerebral infarction	1	0.02

A cross table analysis of ICD-9-CM procedure code versus NIHDI procedure code shows that the code 235082 is mostly used with ICD-9 code 3812; and that codes ICD-9 3832 and 3842 are seldom used. ICD-9 code 3390 is almost never used with NIHDI code 235082 (data not shown).

Conclusion: for both interventions, the primary data selection in MCD 2004 criteria were too large and had to be refined with the use of ICD-9-CM diagnostic codes. The validated algorithm defined by Jacques et al. was thus used for this purpose.[246]

1. CAS:
   The ICD-9-CM procedure code for CAS (3990) is not specific for the carotid vessels and, therefore, had to be combined with the ICD-9-CM diagnostic codes 433.10 'Occlusion and stenosis of carotid artery, without mention of cerebral infarction' or 433.11 'Occlusion and stenosis of carotid artery, with cerebral infarction'

   → algorithm: principal diagnostic 433.10 OR 433.11 AND ICD-9-CM procedure code 3990

2. CEA:
   The ICD-9-CM procedure codes for CEA (3812, 3832, 3842) can also be used for jugular veins, while the NIHDI code can be used to invoice endarterectomies of the vertebral arteries. In addition, this NIHDI code is sometimes applied for the invoicing of CAS. Therefore, the procedure codes need to be associated with the above mentioned diagnostic codes, and the CAS cases have to be excluded.

   → algorithm: principal diagnostic 433.10 OR 433.11

   NOT CAS

   AND NIHDI procedure 235082 OR ICD-9-CM procedure code 3812, 3832, 3842

A total of 3 368 stays are thus selected with these algorithms (after exclusions of 4 stays not classified in the MDC 01 Diseases and Disorders of the nervous system).

Results of procedure definition for CEA/CAS: 3 368 stays selected, of which:

- Carotid Stenting (CAS): 508 stays (15%)

- Carotid Endarterectomy (CEA): 2 860 stays (85%)

Virtually all (98%) of the selected stays are in the APR-DRG 024 "Extracranial vascular procedures" (Table 6.3).

Table 6.3: CEA/CAS: All APR-DRG of selected stays

APR-DRG	Number	Percent
024-EXTRACRANIAL VASCULAR PROCEDURES / 1 – P	3 294	97.80
021-CRANIOTOMY EXCEPT FOR TRAUMA / 1 – P	46	1.37
026-NERVOUS SYST PROC FOR CRANIAL NERV & OTH NERV SYS DISORD / 1 – P	10	0.30
046-NONSPECIFIC CVA & PRECEREBRAL OCCLUSION W/O INFARCT / 1 – M	9	0.27
950-EXTENSIVE PROCEDURE UNRELATED TO PRINCIPAL DIAGNOSIS / 0 – P	5	0.15
952-NONEXTENSIVE PROCEDURE UNRELATED TO PRINCIPAL DIAGNOSIS / 0 – P	2	0.06
023-SPINAL PROCEDURES / 1 – P	1	0.03
045-CVA W INFARCT / 1 – M	1	0.03

6.1.4 Definition of volume

These 3 368 CEA/CAS procedures were distributed across 110 hospitals. The mean number of stays per hospital was 31. See Table 6.4. Two centres are outliers in terms of annual number of interventions i.e. they have more than 130 interventions per year. The proportion of CAS is highly dependent on the centre, and varies from 0 to 100% (Figure 6.2).

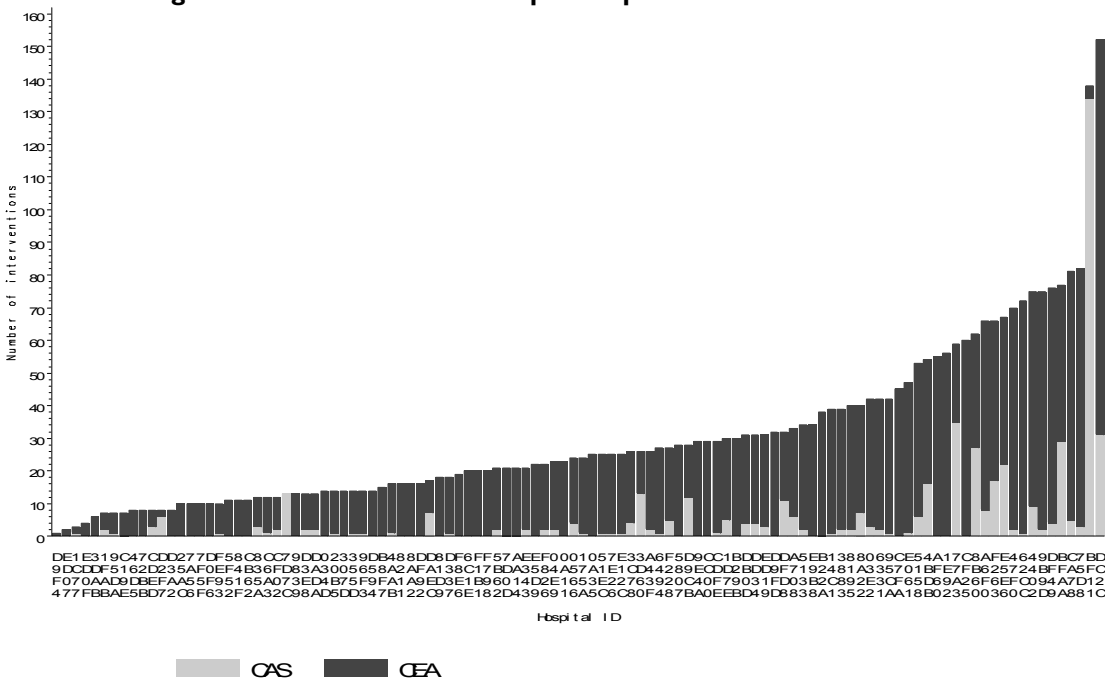
The surgeon volume was available for CEA interventions only since there is no specific NIHDI code for CAS. The average CEA volume per surgeon is 12 per year, performed by 236 surgeons.

Table 6.4: Summary measures of volume per hospital (CEA/CAS) and per surgeon (CEA only)

	Number	Mean	Min	25th Pctl	50th Pctl	75th Pctl	Max	Total
Hospital (CEA/CAS)	110	30.6	1.0	13.0	24.0	39.0	152.0	3368
Hospital (CEA)	109	26.2	1.0	11.0	21.0	34.0	121.0	2860
Surgeon (CEA)*	236	12.0	1.0	3.0	9.0	17.0	60.0	2826*

* information on surgeon is missing for 34 stays

Figure 6.2: CEA/CAS: Volume per hospital



6.1.5 Definition of outcomes

Carotid Endarterectomy (CEA) and Carotid stenting (CAS) both require proficiency since technical errors may lead to abrupt carotid occlusion with stroke or death as possible consequences. Because CEA patients often suffer from diffuse atheromatosis, they also have a higher risk to suffer an acute myocardial infarction in relation to general anaesthesia.

Information on AMI rate and CVA rate was retrieved from the MCD where the secondary diagnosis (as defined in section 4.2.1, page 39) is encoded with the following codes:

- AMI: ICD-9-CM code 410 'Acute myocardial infarction' but after exclusion of patients with AMI as principal diagnosis;

- CVA: ICD-9-CM code 431 'Intracerebral haemorrhage', 432 'Other and unspecified intracranial haemorrhage', 434 'Occlusion of cerebral arteries' or 436 'Acute, but ill-defined, cerebrovascular disease'.

As shown in Table 6.5, in-hospital mortality after CEA/CAS is a very rare event i.e. 0.92%. It was lower after CAS (0.59%) than after CEA (0.98%), although this finding does not permit any conclusion on causality.

Because of possible errors in in-hospital mortality as encoded in the MCD (discussed in section 4.2.4.2 page 41), we retrieved additional information on mortality in the IMA database. This lead to the combined (i.e. combination of MCD and IMA mortality) outcome measure "in-hospital and approximate 30-day mortality" which was 1.27% after CAS/CEA. When this outcome is compared between CAS and CEA, mortality seems higher after CAS (2.0% versus 1.1%). Note that the 30-day mortality is an approximation i.e. minimum 1 day and maximum 60 days, as explained in Chapter 2 (see section 2.3.2.2, page 15).

Acute myocardial infarction (AMI) and stroke (i.e. cerebrovascular accident or CVA) are also rare events after CEA/CAS, i.e., respectively 0.65% and 1.51%.

Table 6.5: CEA/CAS: Outcome measures

		Number Total	**Number** outcome	% outcome
In-hospital mortality	All	3368	31	0.92
	CAS	508	3	0.59
	CEA	2860	28	0.98
In-hospital mortality and approximate 30-day mortality *	All	3304	42	1.27
	CAS	500	10	2.00
	CEA	2804	32	1.14
In-hospital AMI	All	3368	22	0.65
	CAS	508	3	0.59
	CEA	2860	19	0.66
In-hospital CVA	All	3368	51	1.51
	CAS	508	6	1.18
	CEA	2860	45	1.57

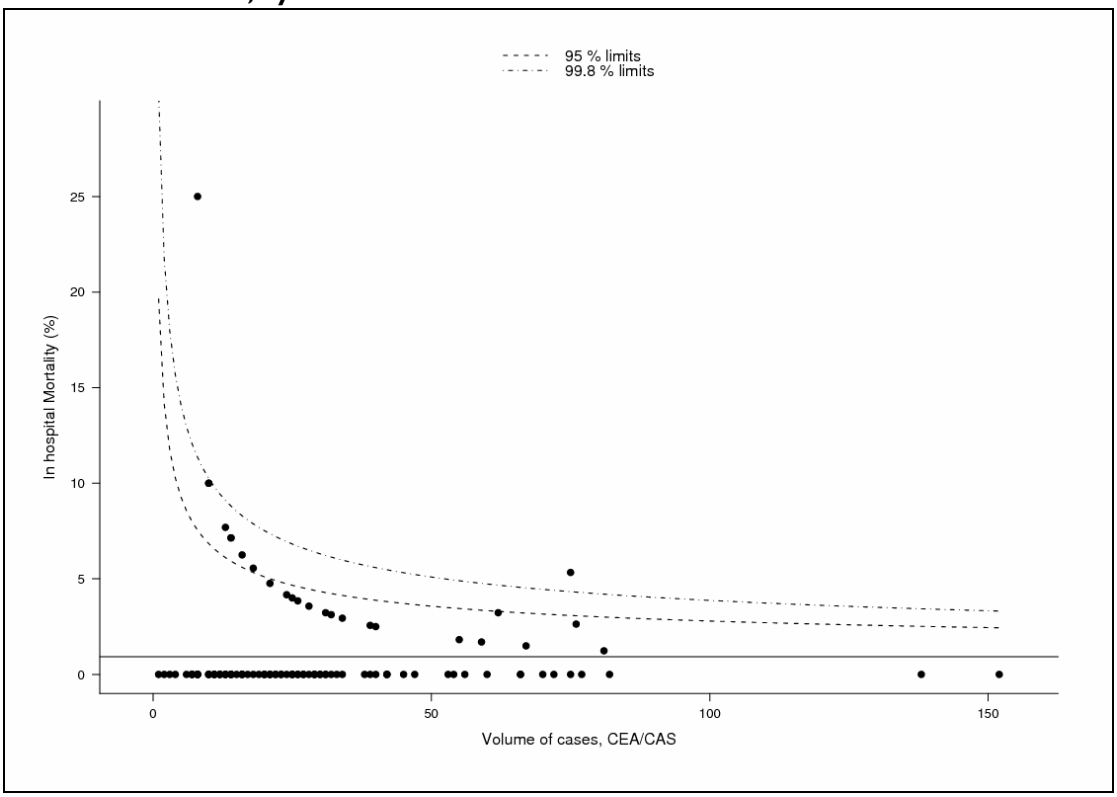
* Only for 3 304 patients with linkage to IMA database

6.1.6 Volume outcome relationship

IN-HOSPITAL MORTALITY RATE, BY HOSPITAL

The funnel plot of the relationship between in-hospital mortality and volume of interventions is presented in Figure 6.3. The horizontal line represents the overall in-hospital mortality i.e. 0.92%. As in-hospital mortality is a very rare event, the majority of the centres have 0% mortality. The funnel plot identifies two outlying centres in terms of outcome: one with low volume and one with medium volume. The two highest volume centres have also 0% mortality.

Figure 6.3: Funnel plot of the in-hospital mortality rate following CEA/CAS treatment, by centre



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

The 110 centres have been divided in three tertiles according to their annual volume of CEA/CAS procedures: hospitals belonging to the first tertile have 16 or less CEA/CAS per year; hospitals in the second tertile perform between 17 and 31 CEA/CAS per year; centres in the third tertile have more than 31 procedures per year. Table 6.6 shows the number of hospitals and stays in each tertile.

The differences in case mix and outcomes between the low volume (first tertile), medium volume (second tertile) and high volume (third tertile) centres are summarized in Table 6.6. Case mix seems very similar in the three tertiles. Observed in-hospital mortality is 1.9% in low volume centres and 0.8% in medium and high volume hospitals.

Table 6.6: CEA/CAS: Differences in case mix and outcomes by differences of volume in centres

		Tertiles based on centre volume			
		N≤16	16<N≤31	N>31	All
Volume					
N cases		411	882	2075	3368
N hospitals		39	36	35	110
Case Mix					
Male	%	64.2	68.1	66.0	66.3
Age	Mean	70.8	71.5	70.9	71.1
	Std	9.0	9.0	9.1	9.0
Charlson index score	Mean	1.3	1.4	1.3	1.4
	Std	1.5	1.6	1.5	1.5
Charlson score ≥ 3	%	17.7	21.2	19.2	19.5
Modified Charlson index score	Mean	1.2	1.3	1.3	1.3
	Std	1.4	1.6	1.5	1.5
Modified Charlson score ≥ 3	%	15.1	19.8	17.4	17.8
Co morbidity: diabetes		20.4	21.2	23.0	22.2
Co morbidity: renal disease		6.6	6.8	8.8	8.0
Severity index (APR-DRG)					
	1	25.8	24.3	28.5	27.1
	2	55.0	58.0	54.7	55.6
	3	17.3	15.3	15.1	15.4
	4	1.9	2.4	1.6	1.9
LOS					
	Mean	8.2	7.8	7.5	7.6
	Median	5.0	5.0	5.0	5.0
	std	9.8	11.0	12.0	11.5
Outcomes					
In-hospital mortality	%	1.9	0.8	0.8	0.9
AMI	%	0.7	0.8	0.6	0.7
CVA	%	1.9	1.9	1.3	1.5

N =number

RELATION BETWEEN HOSPITAL VOLUME AND MORTALITY RATE

Table 6.7 presents the results of the logistic regression. Without adjustment for case mix, an increase of 10% of volume of interventions is associated to a decrease of 4.5% (95% CI -10.5%, 1.3%) in the odds of death. Taking into account the case mix, the estimate of volume effect is rather similar (-5.3%; 95% CI -11.9%, 1.4%). This indicates that case mix is not associated with hospital volume. Age and Charlson score are positively associated with increased in-hospital mortality. A sensitivity analysis with the adjustment based on the modified Charlson score (not including myocardial infarct and cerebrovascular disease in the score) produced similar results. A second sensitivity analysis based only on the volume of CEA interventions also produced a similar estimate of the volume effect i.e. -6% (95% CI -12.4%, 0.5%).

Table 6.7: CEA/CAS: Correlation-corrected logistic regression (GEE) estimates of determinants of in-hospital mortality

Model without adjustment for case mix	Effect[1]	95% CI	
Hospital volume (increase of 10%)	-4.59	-10.51	1.33
Model with adjustment for case mix	**Effect[1]**	**95% CI**	
Hospital volume (increase of 10%)	-5.28	-11.95	1.38
	Odds Ratio	95% CI	
Sex (male vs female)	0.77	0.39	1.51
Age (increase of 1 year)	1.09	1.03	1.17
Charlson score (increase of 1 category)	2.36	1.68	3.32
CAS vs CEA	0.70	0.26	1.88

[1] Effect of 10% increase in volume on the odds of mortality

The association between volume and in-hospital mortality is not robust, as it is sensitive to the one low volume-high mortality centre and the two high volume-low mortality centres, as shown in Table 6.8. When these three centres are excluded, the effect of volume virtually disappears (-1.3%).

Table 6.8: CEA/CAS: Sensitivity analysis (exclusion of 3 centres), Results of logistic regression: relative effect of 10% increase volume on mortality

Model without adjustment for case mix	**Effect[1]**	**95% CI**	
Hospital volume (increase of 10%)	-1.32	-7.75	5.11
Model with adjustment for case mix	**Effect[1]**	**95% CI**	
Hospital volume (increase of 10%)	-1.31	-8.96	6.35

[1] Effect of 10% increase in volume on the odds of mortality

The analyses presented above might be confounded by differences in length of stay between low volume and high volume centres. Therefore, additional sensitivity analyses were performed on the approximate 30-day mortality according to international volume thresholds (for CEA only, used by AHRQ and Leapfrog, see discussion section page 137). As shown in Table 5.12, these analyses confirm the results presented above: one single very high volume centre has 0% mortality, and no differences are observed with respect to centres below or above the AHRQ cut off of 50 CEA per year.

Table 6.9: CEA: in-hospital and 30-day mortality per centre, based on international volume thresholds

Cut off		**Hospital volume category**	**Number centres**	**Number patients**	**Approximate 30-day mortality number**	**%**
AHRQ cut off	50/year	1-50 CEA/year	97	1964	24	1.22
		51-100 CEA/year	11	721	8	1.11
AHRQ and Leapfrog cut off	101/year	≥ 101 CEA/year	1	119	0	0.00

Results based on 2 804 patients with linkage to IMA data

The association between surgeon volume and outcome is not analysed because the number of events (i.e. death) is too low compared to the number of surgeons (31 events for 236 surgeons).

IN-HOSPITAL AMI AND CVA RATE, BY HOSPITAL

The following figures present the AMI rate (Figure 6.4) and CVA rate (Figure 6.5) after CEA/CAS. It was decided not to perform any further analyses on the complication rate because these are rare events, and because complications based on ICD-9 codes may be subject to coding variability.

The latter might explain, for example, why one medium volume centre has a 14% CVA rate after CEA/CAS. Two other centres with lower volume are also outside 99.8% limits of variability.

Figure 6.4: Funnel plot of the AMI rate following CEA/CAS treatment, by centre

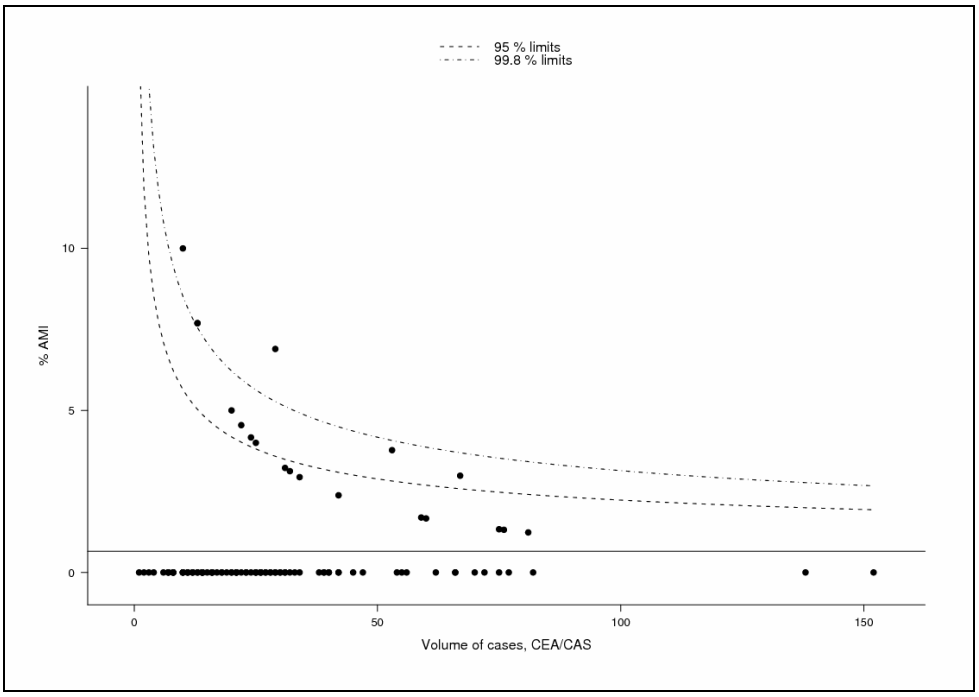
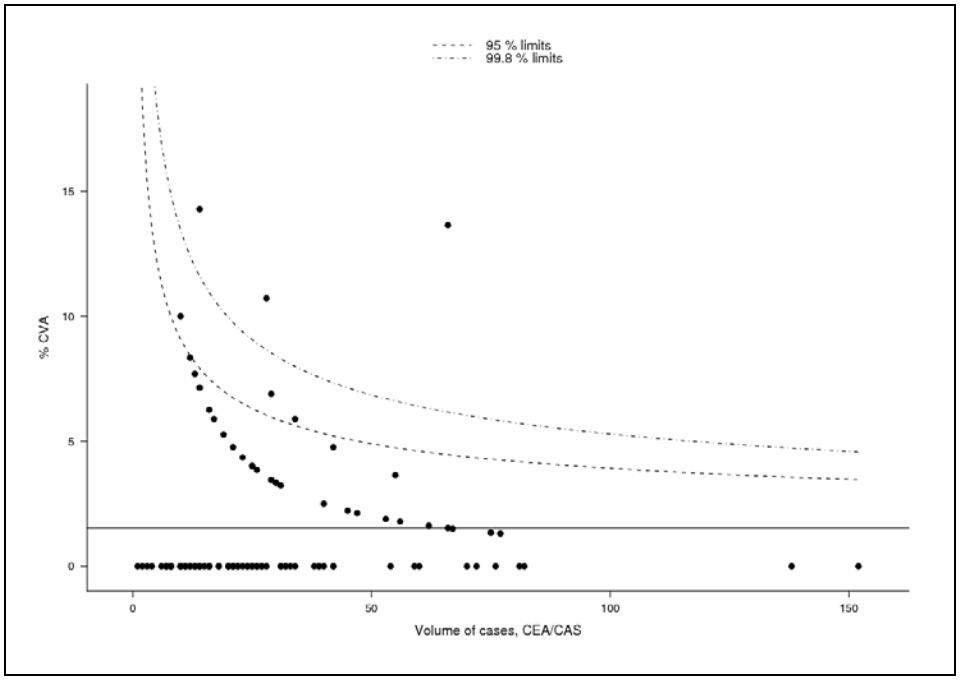


Figure 6.5: Funnel plot of the CVA rate following CEA/CAS treatment, by centre



6.1.7 Discussion

CEA is one of six procedures for which the US Agency for Healthcare Research and Quality (AHRQ) defined a volume indicator (i.e. Inpatient Quality Indicator 7) because outcome may be related to volume.[248] The AHRQ applies a lower threshold of ≥50 procedures per year and an upper threshold of ≥101 CEA per year to benchmark hospitals.[29]

The US Leapfrog group is a non-profit organisation which represents many of the US' largest private and public-sector healthcare purchasers. Together, these purchasers represent more than 34 million Americans.[249] The coalition encourages both patients and payers to select hospitals that meet minimal volume standards for some high-risk procedures.[126] Until March 2004, the Leapfrog group used the cut off of 100 CEA per year to select providers.[93, 115] In April 2004, however, carotid endarterectomy was removed from the Leapfrog Group's list of targeted procedures.[250]

6.1.7.1 External validation of the definition of the procedure

There is no external data source available.

The external experts who commented on this report had their doubts about the number of carotid stentings (CAS) that were selected in the MCD. They argued that the ICD-9-CM procedure code 3990 (see Table 2.1) is probably not commonly used to encode CAS and that some surgeons might use the NIHDI procedure code for carotid arteriography (which was not part of our selection). Because of this limitation, additional analyses were performed on CEA alone.

6.1.7.2 Summarized results of literature review

The systematic literature search identified nine systematic reviews in which the VOA for carotid endarterectomy was studied.[1, 5, 59-65] In all, these systematic reviews identified 38 primary studies of which 12 were published between 2000 and 2005.[93, 115, 116, 251-259] Carotid stenting was not analysed in any of the abovementioned SRs or primary studies.

The discussion below will primarily be based on the meta-analysis by Holt et al. because it is very recent (2007) and of very good quality (see Appendix 7).[61] Holt et al. performed a meta-analysis from 21 articles (885 034 cases) and quantified a pooled effect estimate of hospital volume on stroke and/or mortality rates from CEA. They found that stroke and mortality attributable to CEA occurred less frequently in higher-volume hospitals (odds ratio 0.78 [0.64-0.92]). The critical volume threshold between higher- and lower-volume centres was 79 CEA per year.

This conclusion corresponds to what we concluded on the basis of all systematic reviews in Chapter 2 (see Table 2.2 on page 19); that there is an inverse association between volume (hospital and surgeon) and mortality, as well as stroke rate. A detailed description of this evidence is available in Appendix 9.

6.1.7.3 Comparative analysis of literature and Belgian data

DEFINITION OF VOLUME

As shown in Table 6.4, Belgian hospitals had a mean annual CEA volume of 26 in 2004; 75th Percentile is 34. The application of Holt's hospital volume threshold of 79 CEA per year on Belgian centres would imply that only very few hospitals would comply with the threshold.[61]

In 2004, the Leapfrog Group removed CEA from its list of targeted procedures. This decision was taken on the basis of Birkmeyer's analyses which showed that the absolute difference in mortality between very-low-volume and very-high-volume hospitals was smallest for carotid endarterectomy (1.7 percent vs. 1.5 percent).[115] The analysis by Christian et al. supports the Leapfrog Group's decision: no meaningful volume threshold could be determined for CEA, based on its relationship with mortality.[93] This result made Christian argue that despite the consistent evidence for a relationship between volume and outcome in the literature, it is still not clear how to proceed to minimum procedural volume thresholds which serve as a basis for selective referral to high-volume hospitals.[93]

OUTCOME

In-hospital mortality was 0.98% after CEA (Table 6.5). This is slightly lower than the mean in-hospital mortality of 1.6% (range 0.3 to 5.2%) reported by Holt and associates.[61] The funnel plot (Figure 2.1) revealed two outlying centres with a higher mortality. We would need additional information to investigate and explain these observed results.

The observed in-hospital and approximate 30-day mortality is higher after CAS (2.0%) than after CEA (1.1%) (see Table 6.5). Mas et al. had to prematurely stop their trial when it became clear that the 30-day incidence of any stroke or death was 3.9% after endarterectomy (95% CI, 2.0 to 7.2) and 9.6% after stenting (95% CI, 6.4 to 14.0).[260]

Stroke rate after CEA was 1.51% in Belgian centres (Table 6.5). This is lower than in Holt's meta-analysis which showed a stroke rate of 2.7% (range 0.23 to 6.1%).[61] There is, however, a problem in the coding of post-operative stroke since it is impossible to distinguish between patients who are admitted due to a stroke, and patients who suffer from a stroke post-operatively. Either of these may be miscoded, as Holt points out.[61] Therefore, he suggests that post-operative death is a more valid measure for CEA outcome than stroke.

AMI rate after CEA was 0.66% in Belgium (Table 6.5). This outcome measure has a limited value since the abovementioned coding problem applies in the same way to AMI. In addition, it has to be acknowledged that, in practice, surgeons might miss an AMI because they are not looking for it. Yadav et al., for example, found a 1-year AMI rate of 3.0 and 7.5% after, respectively, CAS and CEA.[261] They defined AMI as an increase of creatinine kinase higher than two times the upper limit with a positive MB fraction.

Christian et al. pointed out that complication rate would actually be a better end point than mortality for procedures such as CEA where death is a rare outcome. The difficulty is, however, the unreliability of the coding of complication data in administrative databases.[93]

PATIENT CASE MIX

Belgian patients with a CEA or CAS procedure were on average 71 years old and predominantly male (66%). One out of five patients (19.5%) had a Charlson score of at least three. Even when AMI and cerebrovascular diseases are excluded from the calculation of the Charlson score (modified Charlson score, explained in section 4.3.2.1), this percentage decreases only slightly to 18%. See Table 6.6.

While age and sex seem comparable with other studies, the percentage of patients with Charlson score $\geq$3, on the other hand, is only 10% in Birkmeyer's study.[115]

Information on the symptomatic status of the patient and the degree of carotid stenosis are supplementary risk factors that are often used for further risk adjustment, but that are not available in the administrative MCD.[61, 115, 256]

VOLUME OUTCOME ASSOCIATION

The meta-analysis performed by Holt et al. showed persistent and reproducible evidence for an inverse relationship between hospital volume and mortality: death attributable to CEA occurred less frequently in higher-volume hospitals (odds ratio 0.76 [0.74-0.81]).[61] The validity of this result was supported by the fact that exclusion of the largest study did not change the pooled effect estimate.

In Belgian data, a CEA volume increase of 10% is associated with a decrease of 6% in the odds of death. The validity of this result is, however, undetermined by the fact that the volume outcome relationship is influenced by three centres (see Table 6.8).

RELATIONSHIP BETWEEN CASE MIX AND VOLUME

Our results (see Table 6.7) show that differences in case-mix between higher- and lower-volume hospitals do not underlie the volume-outcome relationship. This is supported by the analysis by Birkmeyer where significant crude results also remain significant after case mix adjustment.[115]

Key points on volume outcome association for carotid endarterectomy (CEA) and carotid stenting (CAS)

- **In the Minimal Clinical Data 2004, we identified a total of 3 368 hospital stays during which CEA (2 860 stays; 85%) or CAS (508 stays; 15%) was performed. Identification of CAS procedures in MCD is difficult because of lack of a specific procedure code, and, therefore, probably not complete in this study.**
- **A total of 110 hospitals performed CAS/CEA procedures in 2004, with an average of 31 interventions (median 24). The average CEA volume per surgeon is 12 per year, performed by 236 surgeons.**
- **International thresholds for CEA defined by AHRQ and Leapfrog were 50 or 100 procedures per centre per year. In 2004, however, CEA was removed from the Leapfrog list of targeted procedures because the inverse association between volume and mortality was consistent but very weak.**
- **In-hospital mortality (retrieved from MCD) is a very rare event (<1%).The combined outcome measure "in-hospital and approximate 30-day mortality" (the latter is retrieved from IMA database) was 1.3% after CAS/CEA. Regression models were fitted to assess the association between hospital or surgeon volume on this outcome. The following factors were taken into account in all analyses: age, sex, Charlson score (co morbidity) and type of procedure (CEA or CAS).**
- **The literature review indeed concluded that there is an inverse association between volume (hospital and surgeon) and mortality, as well as stroke rate.**
- **Belgian data suggested an observed inverse relationship between in-hospital mortality and hospital volume which is not explained by differences in case mix: 1.9% in low volume centres versus 0.8% in medium and high volume hospitals. However, this relationship is mainly driven by three centres (one low volume-high mortality centre and two high volume-low mortality centres). Associations between volume and approximate 30-day mortality confirm these results.**
- **The association between surgeon volume and outcome was not analysed in Belgian data.**
- **Two complications described in the literature were also studied: AMI and CVA. In-hospital AMI and CVA are rare events, but they suffer from under coding problems.**
- **For CEA, where death is a rare outcome, complication rate (i.e. acute myocardial infarction or stroke rate) would be a better end point. The difficulty, however, is that the coding of these specific complications in the MCD is unreliable.**

6.2 CORONARY ARTERY BYPASS SURGERY (CABG) AND/OR HEART VALVE REPLACEMENT OR REPAIR

6.2.1 General description of procedure

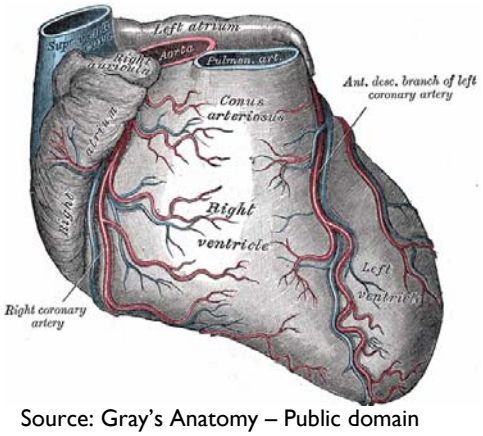
A coronary artery bypass graft (CABG) is a surgical procedure to treat patients with narrowed coronary arteries. Coronary arteries are blood vessels which supply the heart with blood and oxygen (see Figure 6.6). The narrowing (i.e. stenosis) of the coronary arteries is caused by a gradual build up of fatty material (i.e. atheromatous plaque).

The American Heart Association describes CABG as follows: "A coronary artery bypass reroutes the blood supply around a blocked section of the coronary artery. During this procedure, surgeons remove healthy blood vessels from another part of the body, such as a leg or the chest wall. They then surgically attach the vessels to the diseased artery in such a way that the blood can flow around the blocked section."[262]

Most CABG procedures are performed "on-pump": the heart is stopped and blood is pumped by a heart-lung machine. During the past several years, surgeons have started performing "off-pump CABG" (OPCAB) during which the heart continues beating while the bypass graft is sewn in place. According to the American Heart Association scientific statement on bypass surgery, published in 2005, both on-pump and off-pump CABG overall generally result in excellent outcomes and neither type should be judged to be inferior to the other.[263]

CABG is sometimes combined with heart valve replacement surgery (i.e. replacing an abnormal or diseased heart valve with a healthy one) or open heart valvuloplasty (i.e. heart valve repair).

Figure 6.6: The coronary circulation



Source: Gray's Anatomy – Public domain

6.2.2 Primary data selection in MCD 2004

As shown in Table 6.10, the primary selection of data for CABG was based on ICD-9-CM procedure code 361 as proposed by Christian et al., and/or by the NIHDI codes 229585, 229622, 229644 as used by Van Brabandt et al.[57, 93] A total of 8 547 stays are selected. This initial selection does not include the CABG performed in combination with heart valve replacement (i.e. NIHDI procedure code 229526).

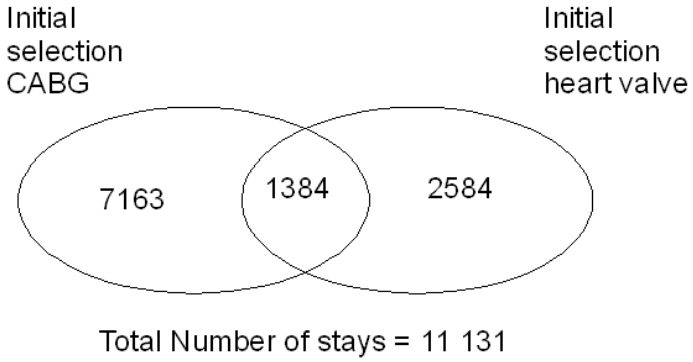
The primary selection of data for heart valve replacement or repair (valvuloplasty) was based on ICD-9-CM procedure codes 351 and 352, or NIHDI procedure codes 229526 and 229600. (see Table 6.10) A total of 3 968 stays are selected.

Because these two procedures can be performed jointly, and because one NIHDI code was missing in the initial selection for CABG (i.e. 229526), it was decided during the definition phase to combine these two primary selections. As 1 384 stays were selected in both primary selection (in the primary selection for CABG and in the primary selection for heart valve replacement), the number of distinct stays in the combined primary selection is 11 131. Figure 6.7 provides an overview of the number of stays in the initial selection.

Table 6.10: CABG and/or heart valve replacement or repair: Primary data selection in MCD 2004

ICD-9-CM procedure code		NIHDI procedure code	
Initial selection of CABG			
361x	Bypass anastomosis for heart revascularization	229585	CABG using two internal mammary arteries. On or off pump.
		229622	CABG using one arterial graft (mammaria, gastroepiploica or an explanted artery) including possible venous bypass(es). On pump.
		229644	CABG using one arterial graft (mammaria, gastroepiploica or an explanted artery) including possible venous bypass(es). Off-pump CABG.
Initial selection of heart valve replacement or repair			
352x	Replacement of heart valve	229526	Surgery on the heart or the great intrathoracic blood vessels that involves the replacement of more than one heart valve, or the replacement of one heart valve in combination with a CABG.
351x	Open heart valvuloplasty without replacement		
		229600	Surgery on the heart or the great intrathoracic blood vessels that involves valvuloplasty or heart valve replacement.
TOTAL STAYS SELECTED= 11 131 (selection 1 OR selection 2)			

Figure 6.7: Number of stays (MCD 2004) from initial selections CABG and/or heart valve repair or replacement



6.2.3 Definition of procedure

A total of 169 stays are excluded because of congenital anomalies (Table 6.11).

Table 6.11: CABG with/without heart valve replacement or repair: Exclusion of congenital anomalies

	Number	Percent
745 Bulbus cordis anomalies and anomalies of cardiac septal closure	30	17.75
746 Other congenital anomalies of heart	129	76.33
747 Other congenital anomalies of circulatory system	8	4.73
759 Other and unspecified congenital anomalies	2	1.18
total	169	

Ninety-eight per cent of the stays (Number = 10 697) are classified in MDC 05 circulatory system. The remaining 2% (Number = 265) are excluded from selection.

As shown in Table 6.12, majority of the selected stays (99%) are in the APR-DRG CABG (165, 166) or in the APR-DRG cardiac valve (162, 163). Since there is no specific APR-DRG for the combination of the two procedures, the selection has to be made based on the NIHDI procedure code.

Table 6.12: CABG with/without heart valve replacement or repair: All APR-DRG from MDC 05

APR-DRG	Number	Percent
166-CORONARY BYPASS W/O MALFUNCTIONING CORONARY BYPASS W/O CARDIAC CATH / 5 – P	**4482**	**41.90**
165-CORONARY BYPASS W/O MALFUNCTIONING CORONARY BYPASS W CARDIAC CATH / 5 - P	**2641**	**24.69**
163-CARDIAC VALVE PROCEDURES W/O CARDIAC CATHETERIZATION / 5 - P	**2577**	**24.09**
162-CARDIAC VALVE PROCEDURES W CARDIAC CATHETERIZATION / 5 – P	**848**	**7.93**
167-OTHER CARDIOTHORACIC PROCEDURES / 5 – P	76	0.71
161-CARDIAC DEFIBRILLATOR IMPLANT / 5 – P	29	0.27
168-MAJOR THORACIC VASCULAR PROCEDURES / 5 – P	18	0.17
160-MAJOR CARDIOTHORACIC REPAIR OF HEART ANOMALY / 5 – P	8	0.07
173-OTHER VASCULAR PROCEDURES / 5 – P	5	0.05
174-PERCUTANEOUS CARDIOVASCULAR PROCEDURES W AMI / 5 – P	3	0.03
177-CARDIAC PACEMAKER & DEFIBRILLATOR REVISION EXCEPT DEVICE REPLACEMENT / 5 – P	3	0.03
171-PERM CARDIAC PACEMAKER IMPLANT W/O AMI, HEART FAILURE OR SHOCK / 5 – P	2	0.02
175-PERCUTANEOUS CARDIOVASCULAR PROCEDURES W/O AMI / 5 – P	2	0.02
180-OTHER CIRCULATORY SYSTEM PROCEDURES / 5 – P	2	0.02
164-CORONARY BYPASS W MALFUNCTIONING CORONARY BYPASS GRAFT / 5 - P	1	0.01
total	10697	100.00

Because of differences in outcomes, the KCE expert group proposed to create three intervention categories:

1. Isolated CABG

Selection based on NIHDI codes 229585, 229622 and 229644 (see Table 6.10 for description of codes).

Because of lack of information in the NIHDI codes, it is not possible to make a distinction between on and off pump CABG.

2. Isolated heart valve replacement or repair

Selection based on NIHDI code 229600 (see Table 6.10). This intervention involves valve repair or valve replacement only.

3. Combination of CABG and heart valve

Selection based on code 229526 (see Table 6.10). This intervention involves valve replacement associated with a CABG.

Table 6.13 shows the repartition of all CABG with/without heart valve repair or replacement in these three intervention categories. In addition, the database contained 17 stays with other combinations of these codes. These stays were excluded from the selection since two different interventions were probably performed during the same stay.

Table 6.13: CABG with/without heart valve replacement or repair: Repartition in three intervention categories

	Number	Percent
Isolated CABG	6813	63.79
Isolated heart valve replacement or repair	1868	17.49
Combination of CABG and heart valve	1655	15.50
no NIHDI code	344	3.22
total	10 680	

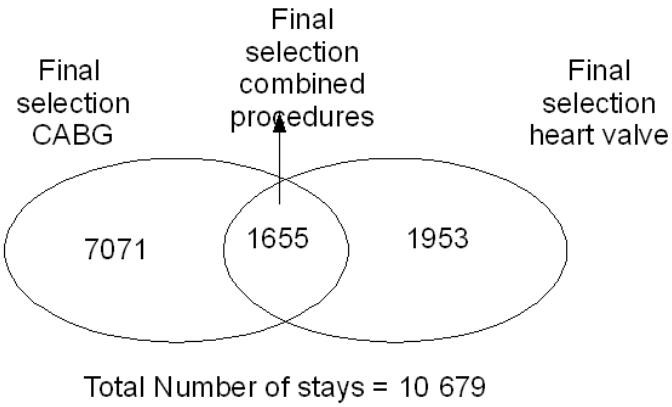
Table 6.13 shows that for 344 of the selected stays (3.2%), none of the selected NIHDI codes was present in the database. These stays were subsequently classified in one of three intervention categories on the basis of their APR-DRG which are listed in Table 6.14.

Table 6.14: CABG with/without heart valve replacement or repair: APR-DRGs from 344 stays without NIHDI code

APR-DRG	Number	Percent
160-MAJOR CARDIOTHORACIC REPAIR OF HEART ANOMALY / 5 - P	1	0.29
162-CARDIAC VALVE PROCEDURES W CARDIAC CATHETERIZATION / 5 - P	24	6.98
163-CARDIAC VALVE PROCEDURES W/O CARDIAC CATHETERIZATION / 5 - P	61	17.73
165-CORONARY BYPASS W/O MALFUNCTIONING CORONARY BYPASS W CARDIAC CATH / 5 - P	126	36.63
166-CORONARY BYPASS W/O MALFUNCTIONING CORONARY BYPASS W/O CARDIAC CATH / 5 - P	132	38.37
total	344	

Finally, this resulted in 10 679 stays. The final repartition of these CABGs with/without heart valve repair or replacement in the three intervention categories is shown in Figure 4.1.

Figure 6.8: Final selection of stays with CABG and/or heart valve repair or replacement



Results of procedure definition for CABG, valvuloplasty or their combination: 10 679 stays selected of which:

- isolated CABG : 7 071 stays;

- CABG combined with heart valve replacement/repair : 1 655 stays;

- isolated heart valve replacement/repair: 1 953 stays.

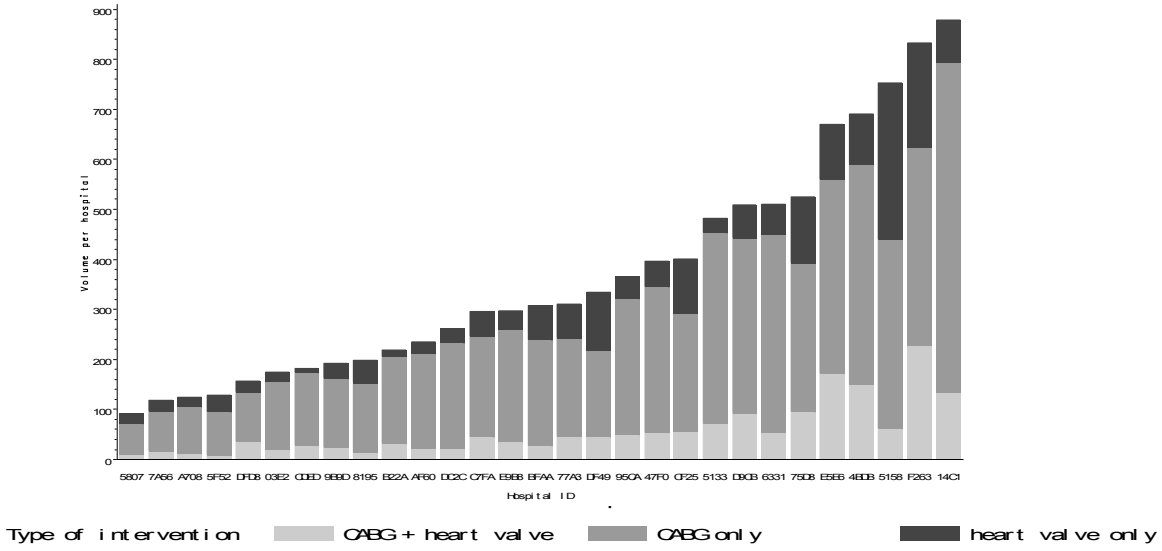
6.2.4 Definition of volume

As is explained in KCE report 14, there are four levels of cardiology services in Belgium:

- A hospitals have no catheterization laboratory (i.e. with diagnostic imaging equipment used to support the catheterization procedure);
- B1 hospitals perform only diagnostic catheterisations (i.e. coronary angiographies);
- B2 hospitals perform both diagnostic and interventional procedures with the exception of CABG;
- B3 hospitals perform all diagnostic and interventional cardiac procedures.57

There are 29 B2/B3 centres. Procedures that were billed in A centres but performed in a B2/B3 centre (with whom the A centre has a transfer agreement) are not included in the following analyses. It concerns 32 CABGs with/without heart valve repair or replacement that were billed in 6 A centres.

Figure 6.9: CABG or/and heart valve replacement/repair: Volume per B2/B3 centre



Summary measures of volume per B2/B3 centre and per surgeon are presented in Table 6.15 and in Table 6.16.

Table 6.15: CABG and/or heart valve replacement/repair: Summary measures of volume per B2/B3 centre

Type of intervention	Number	Mean	Min	25th Pctl	50th Pctl	75th Pctl	Max	Total
CABG + heart valve	29	57.1	8.0	22.0	44.0	62.0	228.0	1655
Isolated CABG	29	242.9	61.0	139.0	211.0	351.0	661.0	7043
isolated heart valve	29	67.2	9.0	23.0	47.0	84.0	313.0	1949
TOTAL VOLUME	**29**	**367.1**	**92.0**	**192.0**	**308.0**	**509.0**	**879.0**	**10647**

Table 6.16: CABG and/or heart valve replacement/repair: Summary measures of volume per surgeon

Type of intervention	Number	Mean	Min	25th Pctl	50th Pctl	75th Pctl	Max	Total
CABG + heart valve	87	19.0	1	6.0	12.0	26.0	125	1655
Isolated CABG	100	68.1	1	25.0	56.5	100.5	291	6806
Isolated heart valve	92	20.2	1	6.0	14.0	25.0	122	1860
TOTAL VOLUME	**104**	**99.2**	**1.0**	**40.0**	**92.0**	**157.0**	**318**	**10321**

6.2.5 Definition of outcomes

Overall, in-hospital mortality was 4.71%, and depends both on the type of operation and on the principal diagnosis (see Table 6.17). In-hospital mortality combined with approximate 30-day mortality was 5.15%. The total number of stays (Number Total) is slightly lower because not all stays could be linked to IMA data.

Table 6.17: CABG and/or heart valve replacement/repair: In-hospital mortality, by type of procedure and principal diagnosis

	Number Total	number Death	% Death
In-hospital mortality	10647	501	4.71
In-hospital mortality and approximate 30-day mortality	10372	534	5.15
In-hospital mortality by type of procedure			
Isolated CABG	7043	249	3.54
Isolated heart valve	1949	109	5.59
CABG + heart valve	1655	143	8.64
In-hospital mortality by Principal Diagnosis			
396 Diseases of mitral and aortic valves	432	25	5.79
410 Acute myocardial infarction	548	70	12.77
414 Other forms of chronic ischemic heart disease	6641	211	3.18
424 Other diseases of endocardium	2215	110	4.97
428 Heart failure	138	17	12.32
441 Aortic aneurysm and dissection	105	13	12.38
996 Complications peculiar to certain specified procedures	180	17	9.44
Other	388	38	9.79

6.2.6 Volume outcome relationship

6.2.6.1 Analysis by hospital

IN-HOSPITAL MORTALITY RATE, BY HOSPITAL

Figure 6.10 presents the observed association between the number of procedures and in-hospital mortality. The horizontal line represents the overall in-hospital mortality i.e. 4.7%. The funnel plot shows that there are a few low volume centres above the expected upper limit of variability, and also that the two highest volume centres are at the lower limit of the expected variability. Data for isolated CABG are presented in Figure 6.11.

Figure 6.10: Funnel plot of the in-hospital mortality of centres after CABG and/or heart valve replacement/repair

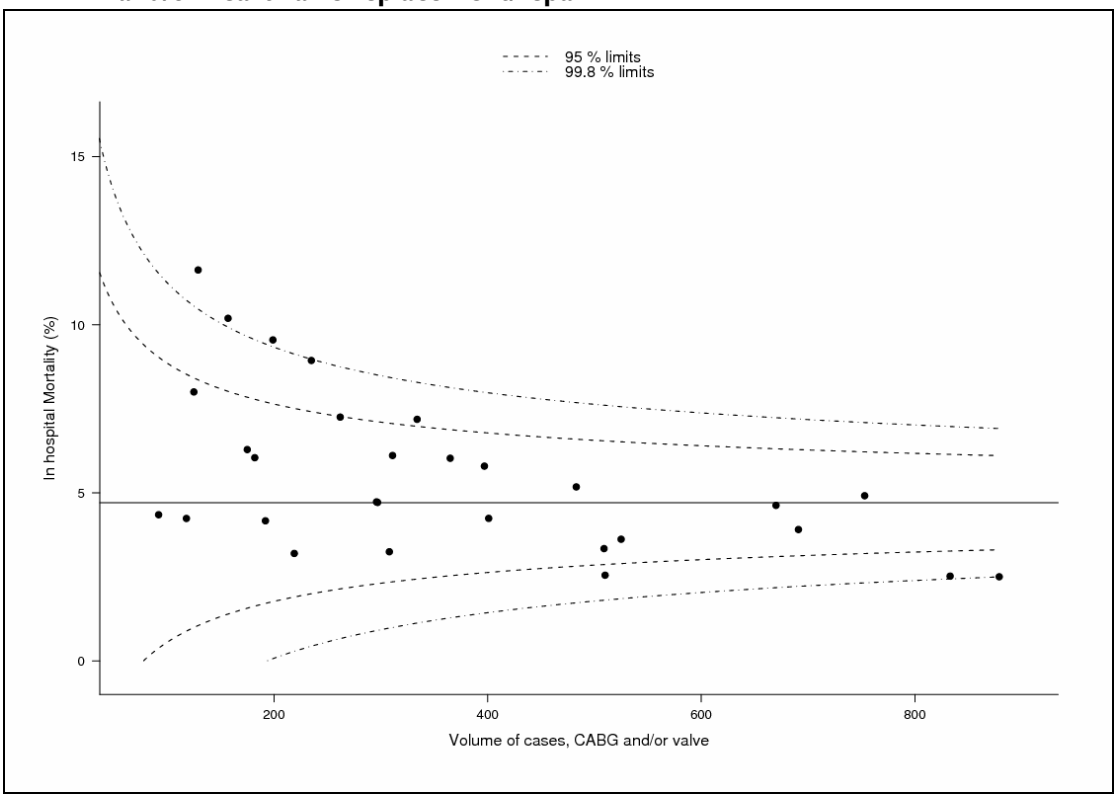
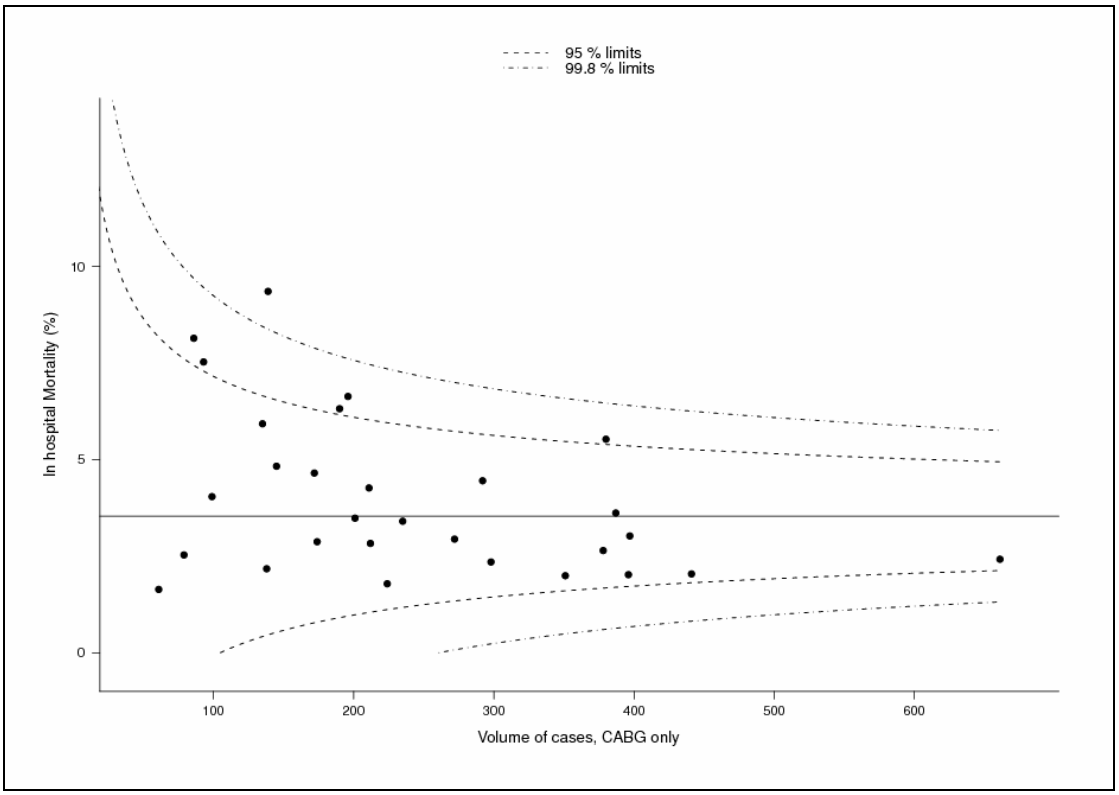


Figure 6.11: Funnel plot of the in-hospital mortality by centre after isolated CABG



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

Table 6.18 shows the differences in case mix by volume of centres. The cut off of volume (219 and 401) divide the 29 centres in three groups of equal size. These thresholds correspond more or less to those used by the Agency for Healthcare Research and Quality (upper threshold of 200 CABG per year) and by the Leapfrog Group (450 CABG per year).[29, 126] More information on these international volume thresholds is provided in the discussion section (page 151).

Age and gender are not associated with volume. There are, however, a higher proportion of patients with a high Charlson score in low and medium volume centres. This is confirmed by a higher proportion of patients with diabetes and renal disease in these centres in comparison to high volume centres.

High volume centres perform more combined operations than low volume centres (18% versus 12%).

Length of stay (LOS) was also longer in low volume centres (17.5 days) than in high volume centres (15.1 days).

In-hospital mortality was inversely associated with the volume of centres: 6.7% in low volume centres, 5.7% in medium volume centres and 3.6% in high volume centres.

Table 6.18: CABG and/or heart valve replacement/repair: Differences in case mix and outcomes by volume of centre

		Tertiles based on volume			
		≤ 219	219-401	> 401	All
Number stays		**1588**	**3206**	**5853**	**10647**
Number centres		**10**	**10**	**9**	**29**
Gender					
Male	%	69.4	69.7	69.3	69.4
Age	Mean	67.3	68.3	67.7	67.8
	Std	11.2	10.2	10.9	10.8
Charlson index score	Mean	1.8	1.8	1.6	1.7
	Std	1.8	1.8	1.7	1.7
Charlson ≥ 3					
	%	30.3	28.1	23.2	25.7
Co morbidities (from Charlson Score)					
Diabetes	%	26.0	24.2	21.8	23.2
Moderate or severe renal disease	%	16.0	16.0	11.6	13.6
Severity index (APR-DRG)					
1	%	2.4	1.4	4.1	3.1
2	%	31.1	22.5	28.5	27.1
3	%	44.1	49.6	43.3	45.3
4	%	22.4	26.5	24.1	24.6
Principal diagnosis					
396 Diseases of mitral and aortic valves	%	3.3	2.5	5.1	4.1
410 Acute myocardial infarction	%	9.3	4.1	4.6	5.1
414 Other forms of chronic ischemic heart disease	%	64.9	63.8	60.9	62.4
424 Other diseases of endocardium	%	15.4	21.4	22.0	20.8
428 Heart failure	%	1.1	0.9	1.6	1.3
441 Aortic aneurysm and dissection	%	1.3	0.9	0.9	1.0
996 Complications peculiar to certain specified procedures	%	1.1	1.1	2.2	1.7
Other	%	3.7	5.2	2.8	3.6
Type of intervention					
CABG + heart valve	%	12.5	12.4	18.1	15.5
Isolated CABG	%	72.4	68.8	63.0	66.2
Isolated heart valve	%	15.1	18.8	18.9	18.3
LOS	Mean	17.5	16.0	15.1	15.7
	Std	12.7	13.1	12.5	12.8
In-hospital mortality	%	**6.7**	**5.7**	**3.6**	**4.7**
In-hospital mortality combined with approximate 30 day mortality	%	**7.3**	**6.3**	**4.0**	**5.1**
Transferred in from another hospital	%	6.9	13.3	22.5	17.4
Transferred out to another hospital	%	7.8	10.0	12.5	11.1

RELATION BETWEEN HOSPITAL VOLUME AND MORTALITY RATE

Results from logistic regression are presented in Table 6.19. Without taking into account differences in case mix, an increase of volume of 10% is associated with a statistically significant (p-value <0.001) change in the odds of death of -4.67% (95% CI -6.8%, -2.6%). After taking into account differences in case mix, this change is smaller but still statistically significant: -3.5% (95%CI -6.5%, -0.5%). A sensitivity analysis adjusting for the modified Charlson score showed similar results.

Table 6.19: CABG and/or heart valve replacement/repair: Correlation-corrected logistic regression (GEE) estimates of determinants of in-hospital mortality

Model without adjustment for case mix	Effect[1]	95% CI	
Volume (increase of 10%)	-4.67	-6.76	-2.58
Model with adjustment for case mix	Effect[1]	95% CI	
Volume (increase of 10%)	-3.51	-6.51	-0.52
	OR	95% CI	
Sex (male vs female)	0.76	0.62	0.93
Age (increase of 1 year)	1.06	1.04	1.07
Charlson score (increase of 1 category)	2.55	2.24	2.91
Principal diagnosis			
396 (Diseases of mitral and aortic valves) vs other	0.44	0.26	0.74
410 (Acute myocardial infarction) vs other	1.41	0.85	2.33
414 (Other forms of chronic ischemic heart disease) vs other	0.40	0.25	0.66
424 (Other diseases of endocardium) vs other	0.33	0.21	0.51
428 (Heart failure) vs other	0.64	0.31	1.30
441 (Aortic aneurysm and dissection) vs other	2.04	1.02	4.06
996 (Complications peculiar to certain specified procedures) vs other	1.16	0.54	2.48
CABG + heart valve vs heart valve only	1.24	0.96	1.59
CABG only vs heart valve only	0.50	0.35	0.70

[1] Effect of 10% increase in volume on the odds of mortality

Sensitivity analyses on the 3 interventions separately show that the effect of hospital volume on in-hospital mortality is less pronounced for isolated heart valve interventions than for procedures which involve CABG (Table 6.20).

Table 6.20: CABG and/or heart valve replacement/repair: Results of logistic regression: relative effect of 10% increase volume on mortality, separate analyses per type of intervention

Intervention	Model with adjustment for case mix	Effect[1]	95% CI	
CABG and heart valve	volume (increase of 10%)	-3.52	-6.51	-0.53
CABG only	volume (increase of 10%)	-3.62	-6.97	-0.27
Heart valve only	volume (increase of 10%)	-1.87	-5.44	0.17

[1] Effect of 10% increase in volume on the odds of mortality

Another sensitivity analysis based on the AHRQ threshold of 200 CABG procedures per year was performed on the Belgian data for isolated CABG.[29] Centres were categorized as low volume (≤ 200 procedures per year, 13 centres) or high volume (> 200 procedures per year, 16 centres). This sensitivity analysis confirms that low volume centres have a higher mortality rate than high volume centres (Table 6.21).

Table 6.21: Sensitivity analysis for isolated CABG, Volume categorized in high volume-low volume centres (using AHRQ cut off)

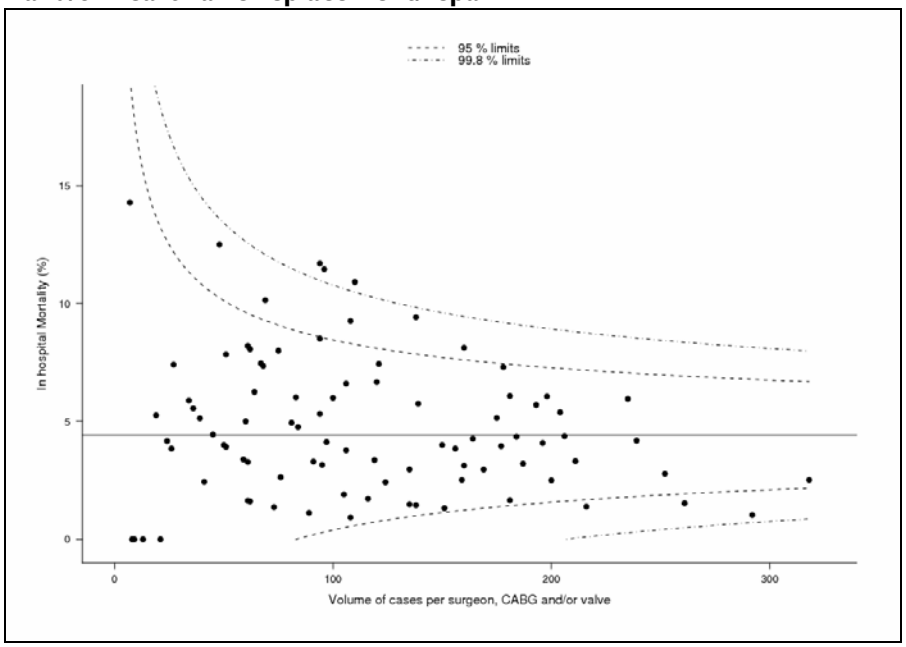
	Number patients	Number death	% death
Low volume (<200 procedures per year)	1707	90	5.3
High volume (≥200 procedures per year)	5336	159	3.0
	Odds Ratio		95%CI
Low versus high volume	1.80	1.34	2.41
Low versus high volume, case mix adjusted	1.77	1.24	2.53

6.2.6.2 Analysis by surgeon

IN-HOSPITAL MORTALITY RATE, BY SURGEON

Analyses at the detailed level of the surgeon are performed. The funnel plot is presented in Figure 6.12.

Figure 6.12: Funnel plot of the in-hospital mortality by surgeon after CABG and/or heart valve replacement/repair



Surgeons having performed less than 5 operations are not displayed in this graphic.

RELATION BETWEEN SURGEON VOLUME AND MORTALITY RATE

Table 6.22 shows how the results from logistic regression confirm the effect of surgeon volume on the risk of death, but to a smaller extent than hospital volume.

Table 6.22: CABG and/or heart valve replacement/repair: Results of logistic regression: relative effect of 10% increase volume on mortality, surgeon volume

Model without adjustment for case mix	Effect[1]	95% CI	
Surgeon volume (increase of 10%)	-2.34	-3.97	-0.71
Model with adjustment for case mix	Effect[1]	95% CI	
Surgeon volume (increase of 10%)	-1.89	-3.64	-0.15
	OR	95% CI	
Sex (male vs female)	0.77	0.62	0.95
Age (increase of 1 year)	1.06	1.05	1.08
Charlson score (increase of 1 category)	2.62	2.34	2.94
396 (Diseases of mitral and aortic valves) vs other	0.47	0.25	0.90
410 (Acute myocardial infarction) vs other	1.26	0.73	2.20
414 (Other forms of chronic ischemic heart disease) vs other	0.45	0.27	0.77
424 (Other diseases of endocardium) vs other	0.36	0.23	0.56
428 (Heart failure) vs other	0.61	0.29	1.28
441 (Aortic aneurysm and dissection) vs other	1.49	0.68	3.25
996 (Complications peculiar to certain specified procedures) vs other	1.10	0.52	2.32
CABG + heart valve vs heart valve only	1.32	1.00	1.73
CABG only vs heart valve only	0.48	0.33	0.70

[1] Effect of 10% increase in volume on the odds of mortality

Analyses which take into account both surgeon as hospital volume concomitantly show that hospital volume plays a more important role than surgeon volume. See Table 6.23.

Table 6.23: CABG and/or heart valve replacement/repair: Results of logistic regression: relative effect of 10% increase volume on mortality, surgeon and hospital volume

Model with adjustment for case mix	Effect[1]	95% CI	
Volume of centre (increase of 10%)	-4.31	-7.11	-1.15
Volume of surgeon (increase of 10%)	-0.07	-2.67	1.14

[1] Effect of 10% increase in volume on the odds of mortality

6.2.7 Discussion

Coronary artery bypass surgery (isolated or in combination with heart valve intervention) is probably one of the most studied procedures in the whole volume outcome literature. It was first mentioned in the seminal article by Luft in 1979 which lead, ten years later, to the first publicly available mortality rates presented per centre and per surgeon in the State of New York.[98, 264] The US Agency for Healthcare Research and Quality (AHRQ) also includes the volume of CABG in its list of quality indicators (Inpatient Quality Indicator 5).[248] The AHRQ applies two thresholds to benchmark hospitals: the lower threshold is ≥100 procedures per year, the upper threshold is ≥200 CABG per year.[29] The US Leapfrog Group uses the cut off of 450 operations per year to select providers.[126]

6.2.7.1 External validation of the definition of the procedure

The external validation of our definition process is based on the data from the Belgian Working Group for Interventional Cardiology (BWGIC) and the Belgian Association for Cardio-Thoracic Surgery (BACTS) for the year 2004, as published in KCE report 66 (table 1.2 on page 11).[265]

These data are introduced by the cardio-thoracic surgeons in the registry of the BWGIC and the BACTS, and unlike the NIHDI data, the BWGIC data are not used to invoice services but for peer reviewed quality control.

The absolute number of CABG corresponds to the selection in our project, but the percentage of CABG combined with heart valve is higher with our definition. There are thus more combined procedures identified in the administrative database than in the registry. This is in contradiction with what was observed in the study of Shahian, which compared definitions of combined CABG-heart valve procedures between clinical and administrative databases.[266]

Table 6.24: Number of CABG and heart valve repairs or replacements: Comparison with BWGIC

	KCE project 2004 Volume Outcome		BWGIC 2004	
	Number	Pct.	Number	Pct.
Isolated CABG	7 071	81%	7 422	85%
CABG & valve	1 655	19%	1 338	15%
Total CABG	8 726	100%	8 760	100%

6.2.7.2 *Summarized results of literature review*

The systematic literature search identified seven systematic reviews (SR) in which the volume outcome association (VOA) for CABG was studied.[1, 5, 59, 60, 62, 64, 66] In all, these systematic reviews identified 29 primary studies of which 10 were published between 2000 and 2004.[93, 115, 116, 267-273] This number was considered sufficient for the discussion. Because the SR by Kalant et al. focuses on CABG, it will be discussed first, followed by the primary studies.[d]

The systematic review by Kalant in 2004 focuses on the relationship between CABG volume and outcome and includes 16 primary studies with CABG performed between 1972 and 1999. The main conclusion of that review is that, although studies from the 70's and the 80's clearly showed that better outcomes were achieved in high volume centres, this relationship has been virtually eliminated in more recent studies. This is explained by the authors by the results of "multifaceted learning curves, improved surgical training and technical advances". Their conclusion is based on all studies comparing low volume centres (usually < 200 interventions per year) to high volume centres (> 200 interventions per year), and showing odds ratio very close to 1 for the recent studies.[66]

On the basis of all seven systematic reviews, it was concluded in Chapter 2 (see Table 2.2 on page 19) that there is an inverse relation between hospital volume and mortality, and between surgeon volume and mortality. It was emphasised, however, that only one out of 7 SRS had a Grade B evidence level. A detailed description of this evidence is available in Appendix 9.

6.2.7.3 *Comparative analysis of literature and Belgian data*

DEFINITION OF VOLUME

The total volume in our study is based on the sum of all isolated CABG, all CABG with intervention on heart valve and all isolated heart valve interventions performed in 2004. A total of 10 673 stays are included in the analysis, distributed over the 29 B2/B3 centres (mean of 367 interventions per centre) and 105 surgeons (mean of 101 interventions per surgeon). The mean annual volume of isolated CABG was 242 interventions per centre and 68 per surgeon. To compare the Belgian CABG volume to the current Leapfrog criterion of 450 CABG interventions per centre per year makes little sense, as this criterion has been heavily criticized.[33, 126] The cut off of 200 procedures is actually considered sufficient to maintain quality according to the American College of Surgeons.[33] Our data show that 16 out of 29 Belgian B2/B3 centres perform more than 200 CABG procedures per year.

[d] The study of Nallamothu is not discussed here, because of the methodological problems described by Kalant.[268]

OUTCOME

In this KCE study, in-hospital mortality was 4.7% overall: 3.5% after isolated CABG, 5.6% after isolated heart valve intervention, and 8.6% after combined intervention. The order of magnitude of these rates can be compared to the in-hospital mortality (or operative mortality, defined as in-hospital or within 30 days after operation) published by the studies summarized in Table 6.25. Differences in rates can be indicative of differences in patient population and procedure selection (isolated CABG or CABG and combined intervention).

Table 6.25: CABG with/without heart valve repairs or replacements: Comparison of Volume-Outcome KCE study with primary studies

Study	Year of interven -tions	Type of procedure	Mortality rate	Age (mean or median)	Male patients (%)
Volume Outcome study KCE	2004	Isolated CABG Isolated heart valve Combination	3.5% 5.6% 8.6%	68 years (for all procedures)	69.4% (for all procedures)
Hannah[271] and Wu[273]	1997-1999	Isolated CABG	2.2%	-	-
Peterson[272]	2000-2001	Isolated CABG	2.7%	66 years	70.7-72.1%
Christian[93]	1999-2000	Isolated CABG	3.9%	65.5 years	71%
Carey[269]	1997-1999	Isolated CABG Isolated heart valve Combination	3.0% 5.0% 9.2%	-	-
Birkmeyer [115, 116]	1994-1999	Isolated CABG	From 4.8% to 6.1%		64.6-65.1%
Gammie[274]	2000-2003	Isolated Heart valve	2.12%		
Ricciardi[275]	2003	CABG (isolated and combined)	3.3%		67.5%

PATIENT CASE MIX

The age of patients (mean age is 68 years old) and 69.4% of them are male. This is comparable to studies described in literature (see Table 6.25).

However, the percent of patients with a Charlson score ≥ 3, as described by Birkmeyer and Ricciardi, does not correspond at all to our data. Birkmeyer and Ricciardi describe that, respectively, 10 and 7% of patients in their study had a Charlson score ≥ 3.[115, 116, 275] This is not at all comparable to the 25% of patients in our study. This might be explained by the fact that in the studies of Birkmeyer, the Charlson Score excluded conditions that "were likely to reflect either the primary indication for surgery or postoperative complications". The authors also explored two alternative approaches to incorporating coexisting conditions in their risk adjustment model, and reported that the three approaches gave virtually identical results. In our analyses, when principal diagnoses were not taken into account in the calculation of the Charlson score, the percent of patients with AMI (ICD-9 410 or 411) was still 22%. This can indicate some overcoding in the co morbidities in the MCD.

No other clinical predictors of CABG mortality, such as those described by Hannan were available in our analysis (e.g. a lower ejection fraction, recent myocardial infarction, left main artery diseases, compromised hemodynamic state, previous open heart operation).[271]

It has to be acknowledged that the Charlson score is not the most appropriate risk score for risk adjustment. For the prediction of mortality after cardiac surgery, the Society of Thoracic Surgeons mortality risk score (STS) and the European System for Cardiac Operative Risk Evaluation (EuroSCORE) scoring system are the two most frequently used risk profile systems.[276]

These scores, however, could not be used in this study because their clinical parameters are not available in the MCD. The STS score is comprised of over 40 clinical parameters, whereas the EuroSCORE involves 18 clinical characteristics.[276, 277] The EuroSCORE system identified three groups of risk factors (with their weights i.e. additive % predicted mortality in brackets):

- Patient-related factors were age over 60 (one per 5 years or part thereof), female (1), chronic pulmonary disease (1), extracardiac arteriopathy (2), neurological dysfunction (2), previous cardiac surgery (3), serum creatinine >200 micromol/l (2), active endocarditis (3) and critical preoperative state (3).
- Cardiac factors were unstable angina on intravenous nitrates (2), reduced left ventricular ejection fraction (30-50%: 1, <30%: 3), recent (<90 days) myocardial infarction (2) and pulmonary systolic pressure >60 mmHg (2).
- Operation-related factors were emergency (2), other than isolated coronary surgery (2), thoracic aorta surgery (3) and surgery for postinfarct septal rupture (4).277

VOLUME OUTCOME ASSOCIATION

Although the 10 studies analyzed in our review all consistently show an effect of centre or surgeon volume on CABG, the effect is always very modest. In the study of Birkmeyer, there is less than 2 absolute percentage points between very low volume and very high volume hospitals.[115] This difference is even smaller in the study of Peterson: 1.1 absolute percentage point.[272] In our study, this difference is 2.3% i.e. from 5.3% in centres with less than 200 CABG per year to 3.0% in centres with at least 200 CABG per year (see Table 6.21).

RELATIONSHIP BETWEEN CASE MIX AND VOLUME

In one study, patient case mix was associated with volume, with low volume centres having higher expected preoperative mortality than high volume centres.[272] This was contradicted by the study of Birkmeyer which showed that case mix of patients was not associated with volume of centres, and that "risk adjustment had negligible effect with respect to CABG procedure".[115, 116]

This is also the case in our study. There is modest evidence that case mix is worse in low volume centres than in high volume centres, as indicated by the higher proportion of patients with high Charlson score, and this despite a higher percentage of combined procedures (with higher mortality) in high volume centres. The adjustment for case mix had thus a modest impact on the effect of volume.

Key points on volume outcome association for coronary artery bypass grafting (CABG) with/without heart valve replacement or repair

- **In the Minimal Clinical Data 2004, we identified a total of 10 679 hospital stays during which CABG with/without heart valve replacement or repair was performed. Three types of procedures are distinguished: 7 071 isolated CABG; 1 655 CABG combined with heart valve replacement/repair; 1 953 isolated heart valve replacement/repair.**
- **These procedures were performed in the 29 so-called B2 and B3 cardiology services. With respect to isolated CABG, these centres had a mean annual volume of 243 CABGs per year (median 211). The average CABG volume per surgeon is 68 per year, performed by 100 surgeons.**
- **The cut off of 200 CABG procedures is actually still considered sufficient to maintain quality according to both the American College of Surgeons as the AHRQ. In 2004, 16 out of 29 Belgian B2/B3 centres reached this volume threshold.**
- **In this study, in-hospital mortality (retrieved from MCD) was 4.7% overall: 3.5% after isolated CABG, 5.6% after isolated heart valve procedure, and 8.6% after combined intervention. In-hospital mortality combined with approximate 30-day mortality was 5.15%. Regression models were fitted to assess the association between hospital or surgeon volume on this outcome. The following factors were taken into account in all analyses: age, sex, Charlson score (co morbidity), principal diagnosis and type of procedure (isolated CABG, isolated heart valve or both combined).**
- **CABG is probably the most studied procedure in the volume outcome relationship, starting in the early 80s. Old studies showed a very consistent and strong effect of volume of centre and surgeon, while results of more recent studies are less straightforward.**
- **Belgian data showed a consistent effect of hospital volume on in-hospital mortality: 5.3% in low volume centres (=<200/year), 3% in high volume centres (> 200/year). This association was not explained by observed differences in case mix. An inverse association was also observed for combined interventions, but not for the isolated heart valve repairs or replacements.**
- **With respect to the association between surgeon volume and in-hospital mortality, the findings of the literature review are confirmed by Belgian data: there is an inverse association between both although the effect is smaller than with hospital volume.**
- **Off pump CABGs are not identifiable in the MCD and can therefore not be taken into account in the case-mix adjustment.**
- **With respect to cardiac surgery, the EuroSCORE (the most frequently used risk profile system in cardiology) would be better for risk adjustment than the applied Charlson score (which takes into account only co morbidities and not patient disease severity). These clinical parameters are, however, not encoded in the MCD.**

6.3 PERCUTANEOUS CORONARY INTERVENTION (PCI)

6.3.1 General description of procedure

Percutaneous coronary intervention (PCI) encompasses a variety of non-surgical procedures used to treat patients with narrowed coronary arteries of the heart.

Initially, the technology was limited to balloon angioplasty which involves advancing a balloon-tipped catheter from an artery in the groin to an area of coronary narrowing.

There, the balloon is inflated whereby it compresses the plaque and widens the narrowed coronary artery so that blood can flow more easily. Catheter and balloon are removed after deflation. Balloon angioplasty is also termed percutaneous transluminal coronary angioplasty (PTCA).[278]

Subsequently, the technique of coronary angioplasty has been expanded by the development of devices that replace or serve as adjuncts to the balloon catheter.

Therefore, PCI has become the generic term for PTCA and other new techniques capable of relieving coronary narrowing such as rotational atherectomy, directional atherectomy, extraction atherectomy, laser angioplasty, implantation of intracoronary stents and other catheter devices.[279]

An intracoronary stent is a wire metal mesh tube used to prop open an artery during angioplasty. The stent is put over a balloon catheter. When the balloon is inflated, the stent expands, locks in place and forms a scaffold. This holds the artery open. The stent stays in the artery permanently and holds it open. Drug-eluting stents (DES) are coated with drugs that are slowly released and help keep the blood vessel from reclosing. Stents that are not coated with drugs are called bare metal stents (BMS).[280]

6.3.2 Primary data selection in MCD 2004

In this report, the term PCI is used instead of PTCA because PTCA only refers to balloon angioplasty, while PCI refers to both angioplasty and stenting.[e]

The primary selection of data was based on ICD-9 procedure code 3601, 3602, and 3605 as proposed by Jollis et al., and/or by the NIHDI codes as used by Van Brabandt et al.[57, 281] A total of 23 037 stays are selected (Table 6.26).

Table 6.26: PCI: Primary data selection in MCD 2004

ICD-9-CM procedure code		NIHDI procedure code	
3601	PTCA single vessel, without thrombolytic agent infusion	589024	Percutaneous endovascular coronary dilatation with/without use of a stent.
3602	PTCA single vessel, with thrombolytic agent infusion		
3605	PTCA multiple vessels		
STAYS SELECTED = 22137		**STAYS SELECTED = 22338**	
TOTAL STAYS SELECTED= 23037 (selection 1 OR selection 2)			

[e] This difference in terminology is not universal, however, as demonstrated by the SR by the German Institute for Quality and Efficiency in Health Care which uses the term PTCA for procedures with and without coronary stent.[67]

6.3.3 Definition of procedure

The majority of the stays (93%) have both the ICD-9 and the NIHDI code (Table 6.27), indicating that there is a good agreement between the two selections.

Table 6.27: PCI: Presence of ICD-9 and NIHDI procedures codes

	Number	Percent	Cumulative Number	Cumulative Percent
ICD-9 only	699	3.03	699	3.03
NIHDI only	900	3.91	1599	6.94
Both	21438	93.06	23037	100.00

Only two percent of the stays are not classified in the MCD 05 circulatory system; these stays are excluded from the selection.

Table 6.28 shows that 73.5% of the selected stays are classified in APR-DRG 175 (PCI without AMI) and 19.9% in APR-DRG 174 (PCI with AMI).

Table 6.28: PCI: All APR-DRGs from MDC 05 Circulatory system

APR-DRG	Number	Percent
175-PERCUTANEOUS CARDIOVASCULAR PROCEDURES W/O AMI / 5 – P	**16 585**	**73.51**
174-PERCUTANEOUS CARDIOVASCULAR PROCEDURES W AMI / 5 – P	**4 497**	**19.93**
192-CARDIAC CATHETERIZATION FOR ISCHEMIC HEART DISEASE / 5 - M	398	1.76
173-OTHER VASCULAR PROCEDURES / 5 – P	292	1.29
165-CORONARY BYPASS W/O MALFUNCTIONING CORONARY BYPASS W CARDIAC CATH / 5 - P	203	0.90
190-CIRCULATORY DISORDERS W AMI / 5 – M	178	0.79
168-MAJOR THORACIC VASCULAR PROCEDURES / 5 – P	67	0.30
198-ATHEROSCLEROSIS / 5 – M	63	0.28
161-CARDIAC DEFIBRILLATOR IMPLANT / 5 – P	58	0.26
191-CARDIAC CATHETERIZATION W CIRC DISORD EXC ISCHEMIC HEART DISEASE / 5 - M	51	0.23
171-PERM CARDIAC PACEMAKER IMPLANT W/O AMI, HEART FAILURE OR SHOCK / 5 - P	34	0.15
202-ANGINA PECTORIS / 5 – M	28	0.12
170-PERMANENT CARDIAC PACEMAKER IMPLANT W AMI, HEART FAILURE OR SHOCK / 5 - P	24	0.11
162-CARDIAC VALVE PROCEDURES W CARDIAC CATHETERIZATION / 5 - P	21	0.09
167-OTHER CARDIOTHORACIC PROCEDURES / 5 – P	18	0.08
166-CORONARY BYPASS W/O MALFUNCTIONING CORONARY BYPASS W/O CARDIAC CATH / 5 - P	11	0.05
950-EXTENSIVE PROCEDURE UNRELATED TO PRINCIPAL DIAGNOSIS / 0 - P	6	0.03
172-AMPUTATION FOR CIRC SYSTEM DISORDER EXCEPT UPPER LIMB & TOE / 5 - P	5	0.02
201-CARDIAC ARRHYTHMIA & CONDUCTION DISORDERS / 5 – M	5	0.02
207-OTHER CIRCULATORY SYSTEM DIAGNOSES / 5 – M	3	0.01
163-CARDIAC VALVE PROCEDURES W/O CARDIAC CATHETERIZATION / 5 - P	2	0.01
169-MAJOR ABDOMINAL VASCULAR PROCEDURES / 5 – P	2	0.01
194-HEART FAILURE / 5 – M	2	0.01
196-CARDIAC ARREST, UNEXPLAINED / 5 – M	2	0.01
160-MAJOR CARDIOTHORACIC REPAIR OF HEART ANOMALY / 5 – P	1	0.00
176-CARDIAC PACEMAKER & DEFIBRILLATOR DEVICE REPLACEMENT / 5 - P	1	0.00
180-OTHER CIRCULATORY SYSTEM PROCEDURES / 5 – P	1	0.00
205-CARDIOMYOPATHY / 5 – M	1	0.00
206-MALFUNCTION, REACTION & COMP OF CARDIAC OR VASC DEVICE OR PROC / 5 – M	1	0.00
951-PROSTATIC PROCEDURE UNRELATED TO PRINCIPAL DIAGNOSIS / 0 - P	1	0.00

Table 6.29 gives an overview of billing codes that are used to invoice drug eluting stents (DES), any stent (DES or BMS) and balloon angioplasty (no stent). These invoice codes are retrievable from the dataset IMPLANT in the MFD.

Table 6.29: PCI: Codes to identify types of stents

Stent or angioplasty	NIHDI implant code
DES	686464
Any stent (DES or BMS)	687886
Balloon angioplasty	687901

Analysis of the codes that were invoiced reveals that 81% of the patients had a stent i.e. 68% BMS and 12% DES. For 10% of the patients the information was missing in the implant dataset. As is explained in the validation section on page 164, this 10% missing data probably refers to DES. Thus, because of the lack of reliable data on the use of DES in the PCI group, the type of stent will not be taken into account in the analyses.

Results of this algorithm for PCI: 22 561 stays selected

6.3.4 Definition of volume

In analogy with KCE report 14, we compared A, B1 and B2/B3 hospitals. A brief description of the four types of cardiology services in Belgium is provided on page 144.

PCI can only be performed in the B2/B3 centres (Number=29). When a patient in need of a PCI is admitted in an A or B1 centre, he/she is transferred to a B2 or B3 centre. If he returns afterwards to the admission centre, the procedure is billed by the A/B1 centre. This was the case for 7% of the PCI, distributed across the 62 A/B1 centres (on average 24 PCI per centre, see Table 6.30). Ideally, PCI procedures which are billed by A or B1 hospital should be added to the volume of B2/B3 hospitals where they are actually performed. In practice, however, information on the B2/B3 hospital was not consistently available in the MCD-MFD data of the referring A or B1 hospital. Therefore, all analyses are limited to patients staying in the B2/B3 centres only (number = 21 028 PCI).

Table 6.30: PCI: Number of stays by type of cardiology program in the admission hospital

Admission hospital with associated cardiology program	Number centres	Number stays	% stays
A hospitals	50	1 080	4.79
B1 hospitals	12	453	2.01
B2/B3 hospitals	29	21 028	93.21
Total		22 561	

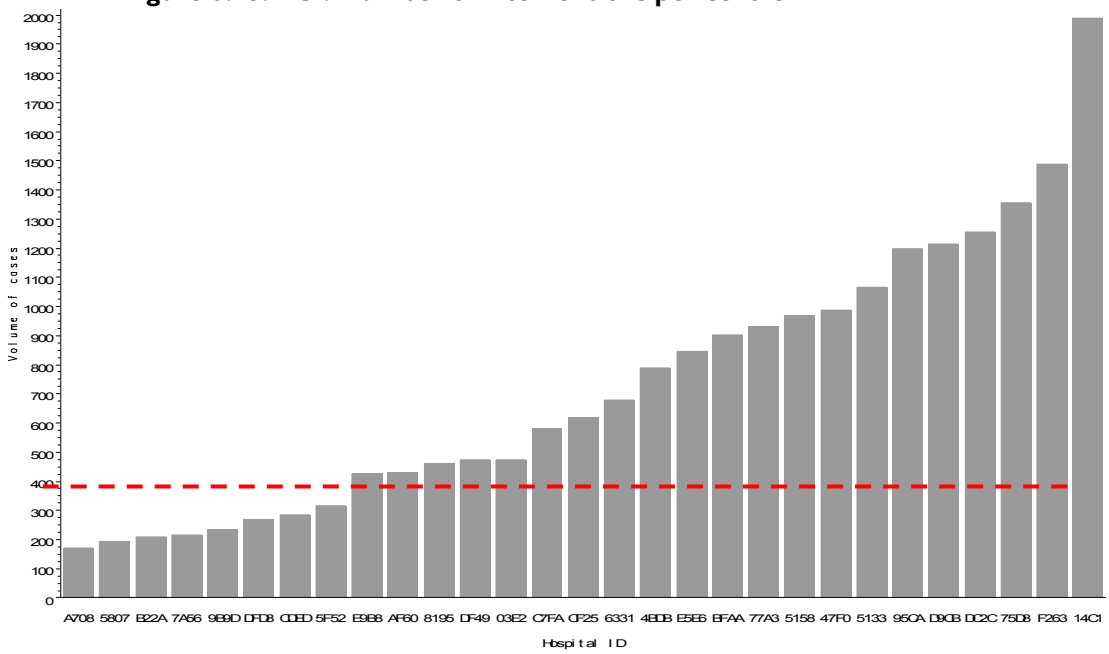
Table 6.31 shows that the average number of PCI per B2/B3 centre is 725 while the median is 620 PCI per annum. Operators perform on average 93 PCI per year, with a median of 51 PCI per annum.

Table 6.31: PCI: Summary measures of volume of stays

	Number	Mean	Minimum	25th Pctl	50th Pctl	75th Pctl	Maximum
Centre	29	725.1	172.0	315.0	620.0	986.0	1989.0
Operator	215	92.9	1.0	6.0	51.0	141.0	852.0

Operator information is available for 19 978 procedures.

Figure 6.13: PCI: Number of interventions per centre



6.3.5 Definition of outcome

The literature review (see Table 2.1 on page 16) identified three procedural complications for PCI: in-hospital death, acute myocardial infarction (AMI), emergency coronary artery bypass graft surgery (CABG). Information on AMI rate was retrieved from the MCD, as described on page 131 for CEA.

Table 6.32 shows the incidence of these outcome measures in the Belgian data. In-hospital mortality after PCI was 1.8%, CABG rate 1.07%, AMI rate 11.6% (after exclusion of stays with AMI as principal diagnosis.

Table 6.32: PCI: Outcome measures

	Number total	number outcome	% outcome
Death	**21 028**	**380**	**1.81**
CABG	21 028	225	1.07
AMI	16 431[a]	1 906	11.60

[a] After exclusion of stays with AMI as principal diagnosis

In-hospital mortality is directly related to the principal diagnosis: the 70% of stays with principal diagnosis 414 (Other forms of chronic ischemic heart disease) have low in-hospital mortality (0.5%) while the 22% of the stays with principal diagnosis 410 AMI have 5.1% in-hospital mortality (Table 6.33).

Table 6.33: PCI: In-hospital mortality in function of principal diagnosis

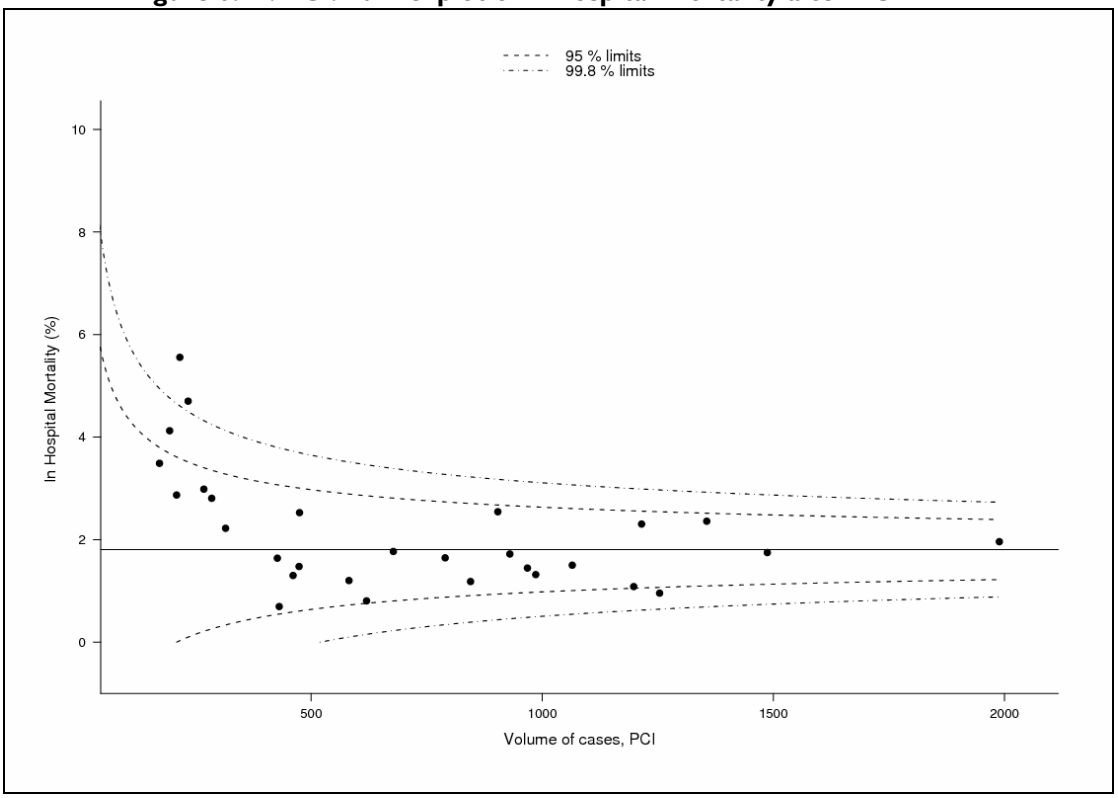
Principal diagnosis	Number of Hospital stays	Number deaths	% death
410 Acute myocardial infarction	**4 597**	**235**	**5.1**
411 Other acute and sub acute forms of ischemic heart disease	539	8	1.5
414 Other forms of chronic ischemic heart disease	**14 815**	**69**	**0.5**
427 Cardiac dysrhythmias	132	21	15.9
428 Heart failure	167	14	8.4
785 Symptoms involving cardiovascular system	20	16	80.0
996 Complications peculiar to certain specified procedures	516	3	0.6
OTHER	242	14	5.8
Total	21 028	380	1.8

6.3.6 Volume outcome relationship

IN-HOSPITAL MORTALITY RATE, BY HOSPITAL

With respect to in-hospital mortality, the funnel plot (Figure 6.14) shows that the majority of the centres are within the expected limits of variability, except for two centres with a low volume which are outside the limits. The horizontal line represents the overall in-hospital mortality i.e. 1.8%. The eight lowest volume centres are above the average mortality.

Figure 6.14: PCI: Funnel plot of in-hospital mortality after PCI



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

Table 6.34 shows the differences in case mix based on the volume of centres (categorized based on tertiles). There are different indicators showing that the case mix is different in low volume centres than in medium or high volume centres. First, the Charlson score is higher (mean and % of stays with score ≥ 3). Secondly, the percentage of stays with AMI as principal diagnosis is higher in low volume centres (28% versus 21% in medium or high volume centres). Thirdly, the APR-DRG severity score is higher in low volume centres (severity 1 in 28% of stays in low volume centres, 40% in high volume centres). There is, however, a difference of 2 years in mean patient's age between the high volume (66 years) and the low volume centres (64 years). In-hospital mortality is 2.8% in centres performing less than 431 PCI/year and around 1.7% in other centres.

Table 6.34: PCI: Case mix in function of volume of hospitals

		Hospital volume			
		< 431	**431-930**	**> 930**	**All**
Number hospitals		10	10	9	29
Number stays		2751	6759	11518	21028
Gender					
Male	%	73.8	73.0	72.8	73.0
Age	Mean	64.2	64.6	66.2	65.5
	Std	12.1	11.5	11.1	11.4
Charlson index score	Mean	1.2	1.0	1.0	1.0
	Std	1.5	1.3	1.4	1.4
Charlson ≥ 3					
	%	15.8	12.3	11.4	12.3
Principal diagnosis					
410 Acute myocardial infarction	%	28.1	20.6	21.1	21.9
411 Other acute and sub acute forms of ischemic heart disease	%	2.5	4.0	1.7	2.6
414 Other forms of chronic ischemic heart disease	%	63.4	71.3	71.6	70.5
427 Cardiac dysrhythmias	%	0.9	0.6	0.6	0.6
428 Heart failure	%	1.2	0.6	0.8	0.8
785 Symptoms involving cardiovascular system	%	0.2	0.1	0.1	0.1
996 Complications peculiar to certain specified procedures	%	1.6	1.5	3.2	2.5
Other	%	2.1	1.3	0.8	1.2
Severity index (APR-DRG)					
1	%	28.3	38.1	40.6	38.2
2	%	51.2	43.2	45.3	45.4
3	%	14.8	14.8	10.5	12.5
4	%	5.7	3.9	3.5	3.9
Length of stay (LOS)	Median	3.0	2.0	2.0	2.0
	Std	7.7	5.8	6.2	6.3
In-hospital mortality	%	**2.8**	**1.6**	**1.7**	**1.8**

RELATION BETWEEN HOSPITAL VOLUME AND MORTALITY RATE

The association between PCI volume and in-hospital mortality is studied in Table 6.35. The volume of centres is statistically significantly associated with in-hospital mortality, as an increase of 10% of the volume is associated with a decrease of 3.8% (-6.1%, -1.5%) of the odds of mortality .The association is entirely explained by the differences in case mix between centres, as no effect remains (estimate 0.3%) after adjustment for patient's age, sex, principal diagnosis and co-morbidities. Sensitivity analyses (with modified Charlson score and without adjustment for Charlson score) showed similar results.

Table 6.35: PCI: Results of logistic regression: relative effect of 10% increase volume on mortality

Model without adjustment for case mix	Effect[1]	95% CI	
Volume (increase of 10%)	-3.81	-6.17	-1.46
Model with adjustment for case mix	Effect[1]	95% CI	
Volume (increase of 10%)	0.30	-2.35	2.94
	OR	95% CI	
Sex (male vs female)	0.98	0.80	1.21
Age (increase of 1 year)	1.06	1.05	1.07
Charlson score (increase of 1 category)	1.99	1.72	2.32
Principal diagnosis			
410 (AMI) vs other	1.54	0.86	2.76
411 (Other acute and sub acute forms of ischemic heart disease) vs other	0.45	0.19	1.04
414 (Other forms of chronic ischemic heart disease) vs other	0.15	0.08	0.27
427 (Cardiac dysrhythmias) vs other	4.20	1.78	9.92
428 (Heart failure) vs other	1.02	0.37	2.79
785 (Symptoms involving cardiovascular system) vs other	> 10	> 10	> 10
996 (Complications peculiar to certain specified procedures) vs other	0.21	0.08	0.61

[1] Effect of 10% increase in volume on the odds of mortality

IN-HOSPITAL CABG AND AMI RATE, BY HOSPITAL

Figure 6.15 how patients had to undergo a CABG after PCI; all centres are within the expected variability. With respect to AMI after PCI, Figure 6.16 shows an extreme variability in rates that can only be explained by radical differences in coding practices. No further analyses are performed on those outcomes.

Figure 6.15: PCI: Funnel plot of CABG after PCI

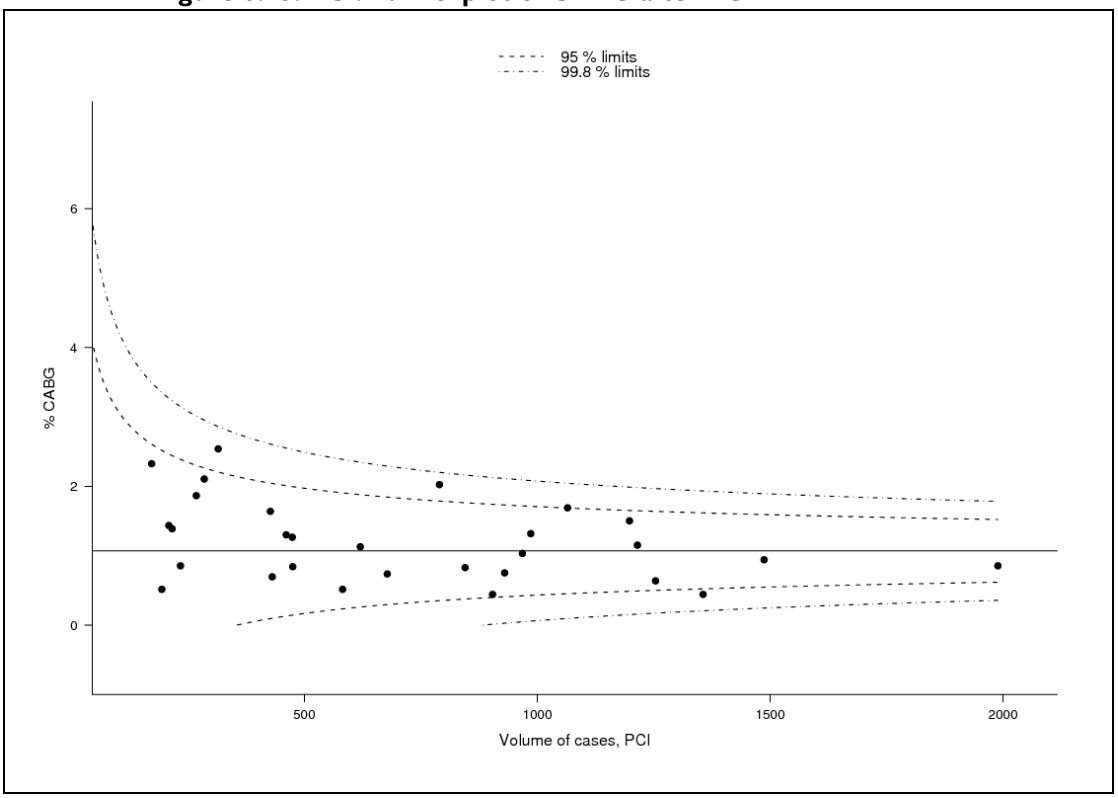
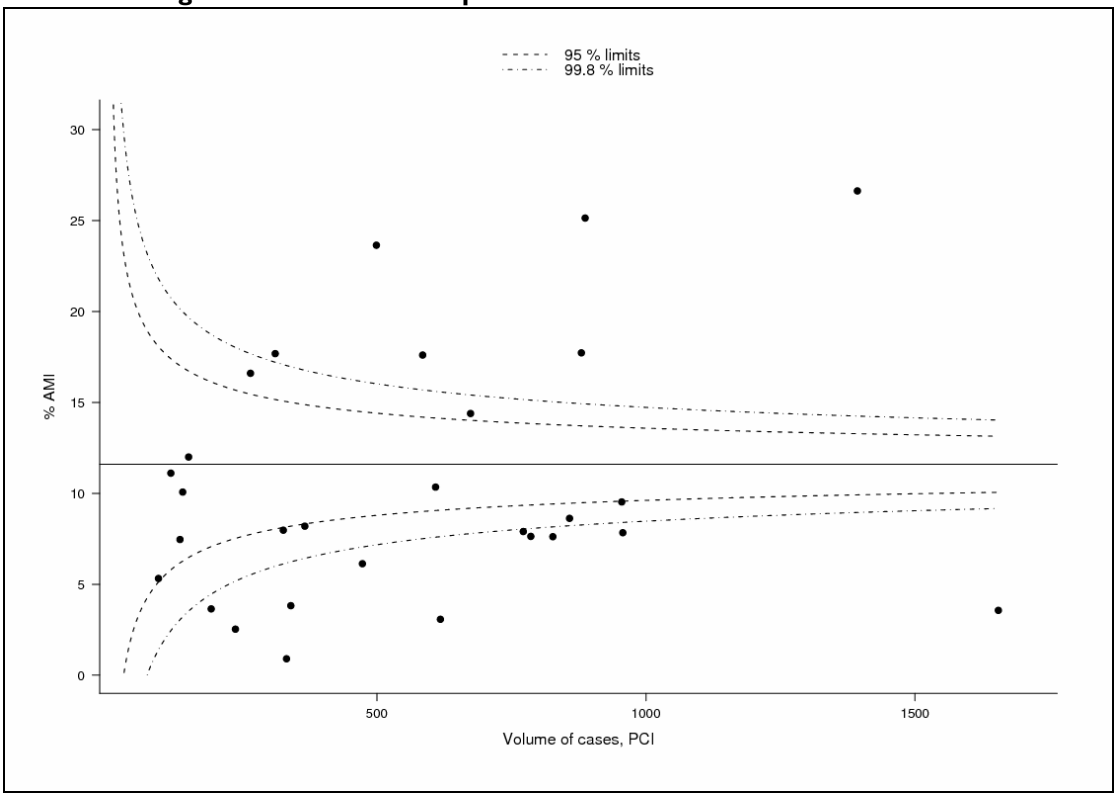


Figure 6.16: PCI: Funnel plot of AMI after PCI



6.3.7 Discussion

PCI is one of those procedures for which volume thresholds have been established. The US Agency for Healthcare Research and Quality (AHRQ) applies a lower threshold of ≥200 PTCA per year and an upper threshold of ≥400 PTCA per year.[29] The US Leapfrog Group uses the cut off of 400 PCI per year to select providers.[126] Finally, the American College of Cardiology/American Heart Association recommends more than 400 PCI per year per hospital and at least 75 PCI per year for operators. Hospitals that perform less than 200 PCI per year are labelled as low-volume hospitals.[279]

6.3.7.1 External validation of the definition of the procedure

The external validation is based on data from the Belgian Working Group for Interventional Cardiology (BWGIC) and the Belgian Association for Cardio-Thoracic Surgery (BACTS) for the year 2004 in Belgium, as published in KCE report 66 (table 1.2 on page 11 in this report).[265]

Table 6.36: PCI: Types of stents

	KCE project 2004 Volume Outcome		BWGIC 2004 [265]	
	Number	Pct.	Number	Pct.
Plain balloon angioplasty	2 144	10%	2 108	9%
Bare Metal Stent (BMS)	15 431	68%	15 696	67%
Drug Eluting Stent (DES)	2 776	12%	5 622	24%
No information	2 210	10%	-	-
Total PCI	22 561	100%	23 426	100%

The total number of PCI is very comparable, but the number of DES is a lot less in our database (i.e. 50%) than registered in the BWGIC register. This could be explained by the reimbursement criteria for DES in Belgium. In 2004, there was only one approved indication for reimbursement i.e. diabetes. Neyt et al. mentioned that DES were also used in non-diabetics (about 14% of non-diabetics received DES or a combination of DES and BMS during their PCI in 2004). In these cases, the hospitals have to bear the additional cost themselves and can not invoice the NIHDI code 686464. Experts that were consulted for KCE report 66, suggested that DES were also implanted in non-diabetics with a high risk of re-stenosis, such as chronic total occlusion, in-stent re-stenosis after prior BMS, multi-vessel stenting, etc. On the other hand, however, Neyt et al. showed that other factors played a role as well. Patients in a private room, for example, had a higher probability of getting a DES.[265]

6.3.7.2 Summarized results of literature review

The systematic literature search identified six systematic reviews in which the VOA for PCI was studied.[1, 5, 59, 60, 64, 67] These systematic reviews discussed 41 primary studies of which 23 were published between 2000 and 2005.[282-304]

The discussion below will primarily be based on the evidence report by the German Institute for Quality and Efficiency in Health Care (IQWiG) because it dates from 2006 and was assessed as very good quality (see Appendix 7).[67] Additionally, several primary studies will be used for the discussion, especially those published in 2004 and 2005 because they allow an evaluation of the volume-outcome relationship in contemporary PCI practice where coronary stents and new anti-platelet agents (i.e. glycoprotein IIb/IIIa receptor blockers) are widespread.

On the basis of all six systematic reviews we concluded in Table 2.2 (page 19) that there were conflicting results in relation to the volume-outcome association between volume (hospital or interventionist) and mortality in case of patients undergoing a PCI for mixed indications (elective and primary) and between hospital volume and emergency CABG rate. The term conflicting results means that there are primary studies that indicate a positive relation with volume and other studies that indicate a negative relation with volume. On the other hand, we concluded that there was an inverse relation between volume (hospital or interventionist) and mortality for primary PCI, and between interventionist volume and emergency CABG rate.

6.3.7.3 *Comparative analysis of literature and Belgian data*

DEFINITION OF VOLUME

Figure 6.13 shows that 8 out of 29 (27%) B2/B3 centres have less than 400 procedures per year (dashed red line in Figure 6.13) which is the upper threshold of several organisations in order to be labelled as high-volume centre (see page 164). Only 1 out of 29 (3%) B2 and B3 centres has less than 200 PCI procedures per year and can be considered a low-volume centre according to the same thresholds.

Operators perform on average 93 PCI per year, with a median of 50 PCI per annum, as is illustrated in Table 6.31. This means that many of the Belgian operators do not comply with the operator threshold of 75 PCI per year by the American College of Cardiology.[279]

OUTCOME

Overall in-hospital mortality after PCI was 1.81% in Belgian B2/B3 hospitals (Table 6.32). There is, however, a relation with the principal diagnosis: in-hospital mortality is much higher among patients admitted with a myocardial infarction i.e. 5.11% (Table 6.33). Similar differences are present in scientific literature where a distinction is made between articles that only study primary PCI and those that analyse PCI for all indications (i.e. primary and elective PCI). Primary PCI is synonym for immediate PCI and is performed on patients with an acute myocardial infarction, within a limited number of hours after the onset of symptoms. Epstein et al. analysed more than half a million hospitalizations with a PCI (primary and elective) performed nationwide in the USA between 1998 and 2001. They found a crude in-hospital mortality of 1.50% for these patients; this seems comparable with the Belgian 1.8%.[283] Zahn et al. used data of almost 5 000 patients treated with a primary PCI in German hospitals between 1994 and 2000. In-hospital mortality was 9.3% for these patients with an acute myocardial infarction.[287] This is even higher than the 5.11% that was found in Belgian data.

Table 6.32 shows a CABG rate of 1.07% in Belgian hospitals after PCI. This percentage is comparable with the one mentioned by Hannan et al.; same-stay CABG rate of 0.91% after PCI performed in New York between 1998 and 2000.[284] Moscucci et al., on the other hand, observed an emergency CABG rate of 0.44% among 18 500 PCIs which were performed in Michigan hospitals in 2002.[285]

AMI rate after PCI is extremely high in Belgian data: 11.6% after exclusion of stays with AMI as principal diagnosis (see Table 6.32). Moscucci et al. found a MI rate of only 1.37% in their database.[285] This discrepancy proves once again that the use of codes from administrative data is not an ideal method to accurately identify the occurrence of AMI as a complication, as was also the case for CEA.

PATIENT CASE MIX

Almost three out of four (73%) Belgian PCI patients are male. They have an average age of 65.5 and a mean Charlson score of 1.0 (Table 6.34). In relation to age and gender, Belgian patients are very comparable to those described in the primary studies that were identified. There were, however, only a few studies that mention the Charlson score. Kansagra et al. found a mean Charlson score of 0.9 among almost 100 000 patients who underwent PCI in the USA in 2001.[290] Ho et al., on the other hand, found much lower Charlson scores, but they described an evolution in mean Charlson score across different time periods: 0.24 in 1984-1987, 0.35 in 1988-1992 and 0.46 in 1993-1996.[301] These numbers suggest that patients who received PCI become frailer over time, which could explain the mean score of 1.0 in Belgian data.

Table 6.34 shows a difference in case mix between Belgian centres i.e. a higher percentage of patients with Charlson score $\geq$3, with AMI as principal diagnosis or with a higher severity index in low-volume hospitals. When we compare these results with published studies that are also limited to administrative data, there is a resemblance. Epstein et al. also found a slightly higher proportion of PCI patients with a myocardial infarction in low-volume (<400 PCI/year) hospitals. The prevalence of co-morbid conditions, however, was generally comparable between patients of low and high volume hospitals.

There was no information on severity score.[283] McGrath et al. described how lower-volume hospitals were more likely to perform PCI on patients admitted with AMI, but, on the other hand, there were more patients with a Charlson score >1 in high-volume hospitals. Age was similar across volume categories.[304] The external experts who commented on our study were not surprised by the higher percentage of patients with AMI in low-volume centres. They argued that low-volume centres perform more emergency PCIs. A patient who presents himself with an acute myocardial infarction at a low-volume hospital can not wait to be transferred to a high-volume hospital, whereas patients who are scheduled for an elective PCI have a choice.

However, many studies use clinical data to adjust for case mix. Hannan et al. found that "patients in higher-volume hospitals were significantly older and were significantly more likely to have had left main disease, more vessels diseased, a type C lesion, previous stroke, aortoiliac disease, congestive heart failure, previous PCIs and open heart operations, and stent during the index procedure." Patients in lower-volume hospitals were significantly more likely to have had chronic obstructive pulmonary disease and diabetes. Nevertheless, Hannan emphasised that these statistically significant differences were more likely to be attributable to large sample sizes rather than to clinical significance.[284]

Overall, the studies based on clinical databases use a variety of clinical characteristics: [284, 285, 297-299]

- Cardiac risk factors: smoking, hypertension, hypercholesterolemia, diabetes, mellitus, family history of coronary artery disease;
- Medical history of: angina, congestive heart failure, chronic obstructive pulmonary disease, heart surgery, gastro-intestinal bleeding, extra-cardiac vascular disease, atrial fibrillation, renal failure requiring dialysis, stroke, myocardial infarction, CABG, PCI;
- Clinical presentation at time of PCI: left ventricular ejection fraction, creatinine level, myocardial infarction within 24 h, rescue PCI, unstable angina, cardiogenic shock, cardiac arrest, anemia (haemoglobin <10g/dl), heart rate, blood pressure, Killip class which measures the severity of heart failure with myocardial infarction;[305]
- Procedural characteristics: time to treatment, extent of coronary heart disease (one-, two- or three-vessel disease), visible thrombus on the initial coronary angiogram, type C lesion, severe calcification, stent, glycoprotein IIb/IIIA blockers, total contrast/case (ml), vasopressor treatment, intra-aortic balloon pump, angiographic success.

The Charlson score gives no information on the cardiac status of the patient and is therefore not really appropriate to assess early outcomes following PCI.

VOLUME OUTCOME ASSOCIATION

After adjustment for case mix, based on gender, age, Charlson score and principal diagnosis, the association between hospital volume and mortality in Belgian data 2004 disappeared. This result corresponds with the conflicting results in relation to the volume-outcome association between volume (hospital or interventionist) and mortality in case of patients undergoing a PCI for mixed indications (elective and primary) in systematic reviews (see Table 2.2 on page 19).

For primary PCI, however, the systematic reviews concluded that there was an inverse relation between volume (hospital or interventionist) and mortality. Unfortunately, the administrative data did not permit a sub-analysis of primary PCI.

Key points on volume outcome association for Percutaneous Coronary Intervention (PCI)

- **In the Minimal Clinical Data 2004, we identified a total of 21 028 hospital stays during which PCI was performed in one of 29 B2 or B3 centres. Due to the lack of reliable data on the use of drug eluting stents (DES) during these stays, the type of stent (i.e. DES, bare metal stent or balloon angioplasty) could not be taken into account in the analyses**
- **The 29 B2/B3 centres had an average volume of 725 PCIs per year (median 620). Eight out of 29 centres (27%) applied to the international cut off of 400 PCIs per year. The average PCI volume per operator is 93 per year, performed by 215 physicians.**
- **Overall in-hospital mortality after PCI was 1.81% in Belgian B2/B3 hospitals, but is much higher among patients who are admitted with a myocardial infarction i.e. 5.11%. Regression models were fitted to assess the association between hospital or interventionist volume on this outcome. The following factors were taken into account in all analyses: age, sex, Charlson score (co morbidity) and principal diagnosis.**
- **The literature review could only conclude on an inverse relation between volume and mortality for primary PCI, which is synonym for immediate PCI (performed on patients with an acute myocardial infarction, within a limited number of hours after the onset of symptoms).**
- **Belgian data show that observed hospital volume is statistically significantly associated with in-hospital mortality which is 2.8% in low volume centres (< 431 PCI) versus 1.6 and 1.7% in medium and high volume centres. However, this association is entirely explained by the differences in case mix between centres, as no effect remains after risk adjustment. Low volume hospitals had a higher percentage of patients with Charlson score ≥ 3, with AMI as principal diagnosis or with a higher severity index. A subset analysis on primary PCI patients was not possible in Belgian data.**
- **The association between operator volume and outcome was not analysed in Belgian data.**
- **Three complications described in the literature were also studied: emergency CABG (1.07%), AMI (11.6%). However, AMI is not a reliable indicator since huge coding differences exist between the centres.**
- **Some procedures are not identifiable in the MCD and can therefore not be taken into account in the case-mix adjustment. This limitation applies to the type of stent implanted during PCI (i.e. drug eluting stent, bare metal stent or balloon angioplasty).**
- **With respect to PCI, the EuroSCORE (the most frequently used risk profile system in cardiology) would be better for risk adjustment than the applied Charlson score (which takes into account only co morbidities and not patient's cardiac status). These clinical parameters are, however, not encoded in the MCD.**

7 RESULTS FOR THREE ORTHOPAEDIC PROCEDURES

7.1 ELECTIVE TOTAL HIP REPLACEMENT/ARTHROPLASTY (THR)

7.1.1 General description of procedure

Hip replacement or arthroplasty is a procedure in which damaged or diseased parts of the hip joint are removed and replaced with an artificial joint, called prosthesis. There are three types of hip replacement surgery:

- Partial hip replacement: either the femoral head or the hip socket (acetabulum) will be replaced with a prosthetic device;
- Total hip replacement (THR): both femoral head and acetabulum are replaced with a prosthesis;
- Revision hip replacement: a hip prosthesis that no longer fits or functions well is replaced.

This report only studies the volume-outcome relationship for the primary (original) total hip replacement.

7.1.2 Primary data selection in MCD 2004

First selection of data was based on ICD-9-CM procedure and NIHDI procedure codes as proposed by Taylor et al. and by Jacques et al. This selection resulted in a total of 22 369 stays (Table 7.1).[246, 309]

Table 7.1: Total hip arthroplasty: Primary data selection in MCD 2004

SELECTION 1 ICD-9-CM procedure code		**SELECTION 2 NIHDI procedure code**	
8151	Total hip replacement	289041	Hip arthroplasty with a femoral prosthesis
8152	Partial hip replacement		
8153	Revise hip replacement	289063	Hip arthroplasty with an acetabulum prosthesis
		289085	Hip arthroplasty with a total prosthesis (acetabulum and femoral head)
		293440	Removal of a total hip prosthesis and placement of a new total prosthesis
STAYS SELECTED =22 048		**STAYS SELECTED = 20 267**	
TOTAL STAYS SELECTED=22 369 (selection 1 OR selection 2)			

7.1.3 Definition of procedure

Table 7.2 shows that for 89% of all stays both ICD-9-CM procedure and NIHDI procedure were encoded. Nine percent of the stays (number = 2 102) did not have a NIHDI procedure code for hip replacement (or revision) on their bill.

Table 7.2: Total hip arthroplasty: Cross table ICD-9 and NIHDI procedures codes

	Number	**Percent**	**Cumulative Number**	**Cumulative Percent**
ICD-9 only	2 102	9.40	2 102	9.40
NIHDI only	321	1.44	2 423	10.83
Both	19 946	89.17	22 369	100.00

Table 7.3 shows which NIHDI procedures were billed for the 2 102 stays with only an ICD-9-CM procedure code for THR. Forty-seven percent of these stays concerned the surgical treatment of a femoral neck fracture with prosthesis. This is a typical non-elective, emergency procedure. Finally, all 2 102 stays were excluded from further analysis since they do not belong to the target population of patients with elective hip replacement. To avoid flaws in the data selection, stays without ICD-9-CM codes were also not considered (number = 321).

Table 7.3 Total hip arthroplasty: first NIHDI procedure codes for stays with only an ICD-9-CM procedure code for THR (number = 2 102)

NIHDI code		COUNT	PERCENT
289402	Surgical treatment of a femoral neck fracture with prosthesis.	1 545	46.78
294803	Supplementary fee for continuous traction through a pin to manage fractures of the lower limbs.	334	10.11
299386	Continuous skin traction of a lower limb.	157	4.75
290286	Femoro-tibial arthroplasty with articulated prosthesis	121	3.66
289365	Surgical treatment of a per- or intertrochanteric femoral fracture.	89	2.69
298745	Non-surgical treatment of a femoral neck fracture (without reduction).	75	2.27
298023	Non-surgical treatment of a hip luxation.	71	2.15
289026	Hip prosthesis with interposition of tissue or a cup.	65	1.97

Among the remaining 19 946 stays, there were 14 636 stays with one THR which were encoded with the NIHDI procedure code 289085 as well as the ICD-9-CM procedure code 8151. In order to obtain a homogeneous population, the following procedures were excluded from further analysis: partial arthroplasties (number = 3246), revision hip replacements (number = 2117) and bilateral procedures (number = 133). In addition, some data abnormalities were also discarded (number = 27).

After this first selection on the basis of procedure code, we looked at the available clinical information. Major Diagnostic Category (MDC) 08 "Diseases & disorders of the musculoskeletal system" represented 99.16% of the 14 636 stays. Their APR-DRGs are presented in Table 7.4.

Table 7.4 Total hip arthroplasty: All APR-DRGs of stays from MDC 08-Diseases & disorders of the musculoskeletal system AND with one hip arthroplasty coded in NIHDI procedure AND in ICD-9-CM procedure

	COUNT	PERCENT
302-MAJOR JOINT & LIMB REATTACH PROC OF LOWER EXTREM EXC FOR TRAUMA / 8 P	13 204	90.98
301-MAJOR JOINT & LIMB REATTACH PROC OF LOWER EXTREMITY FOR TRAUMA / 8 P	1 302	8.97
300-BILATERAL & MULTIPLE MAJOR JOINT PROCS OF LOWER EXTREMITY / 8 - P	7	0.05
TOTAL	14 513	100.0

The hip replacements after trauma i.e. APR-DRG 301 were excluded (9% of the MDC 08) since this study focuses on elective THR. A remaining handful of stays in APR-DRG 300 with bilateral and multiple procedures were also excluded.

Table 7.5 shows that the majority (90.8%) of the stays in APR-DRG 302 had principal diagnosis 715 "Osteoarthrosis & allied disorders". 66 stays with principal diagnosis 996x "Complications peculiar to certain specified procedures" and 3 stays with V54x "Other orthopaedic aftercare" were rejected.

Table 7.5 Total hip arthroplasty: List of all principal diagnostic ICD-9 codes in APR-DRG 302

Principal diagnosis	Freq	%
715 Osteoarthrosis & allied disorders	11988	90.79
733 Other disorders of bone & cartilage	922	6.98
996 Complications peculiar to certain specified procedures	66	0.50
716 Other & unspecified arthropathies	51	0.39
730 Osteomyelitis, periostitis & other infections involving bone	35	0.27
755 Other congenital anomalies of limbs	21	0.16
718 Other derangement of joint	20	0.15
714 Rheumatoid arthritis & other inflammatory polyarthropathies	17	0.13
736 Other acquired deformities of limbs	16	0.12
198 Secondary malignant neoplasm of other specified sites	11	0.08
719 Other & unspecified disorders of joint	10	0.08
754 Certain congenital musculoskeletal deformities	8	0.06
732 Osteochondropathies	6	0.05
905 Late effects of musculoskeletal & connective tissue injuries	4	0.03
711 Arthropathy associated with infections	3	0.02
720 Ankylosing spondylitis & other inflammatory spondylopathies	3	0.02
724 Other & unspecified disorders of back	3	0.02
813 Fracture of radius & ulna	3	0.02
V54 Other orthopaedic aftercare	3	0.02
171 Malignant neoplasm of connective & other soft tissue	2	0.02
727 Other disorders of synovium, tendon, & bursa	2	0.02
812 Fracture of humerus	2	0.02
170 Malignant neoplasm of bone & articular cartilage	1	0.01
277 Other & unspecified disorders of metabolism	1	0.01
712 Crystal arthropathies	1	0.01
721 Spondylosis & allied disorders	1	0.01
722 Intervertebral disc disorders	1	0.01
723 Other disorders of cervical region	1	0.01
731 Osteitis deformans & osteopathies associated with other disorders classified elsewhere	1	0.01
756 Other congenital musculoskeletal anomalies	1	0.01
TOTAL	13204	100.00

The final step in the selection was to discard 1 239 patients who had undergone two THR during two consecutive hospitalizations in the period 2004-2005. This was done to avoid confusion in identifying and allocating adverse events. The final population comprised only patients with one THR in 2004-2005. The volume per hospital was thus derived from this selection of stays.

Results of definition: 11 856 stays with one elective total hip replacement were selected (for 11 856 patients).

7.1.4 Definition of volume

Table 7.6 shows that 115 hospitals performed 103 elective total hip arthroplasties on average in 2004. One hospital performed 608 procedures, the second in volume accounting for 396 procedures. Three quarters of the hospitals performed 121 elective THR or less.

Table 7.6: Total hip arthroplasty: Summary measures of volume per hospital

Number	Minimum	Mean	25th Pctl	50th Pctl	75th Pctl	Maximum
115	5.0	103.1	42.0	84.0	121.0	608.0

Figure 7.1: Total hip arthroplasty: Volume per hospital

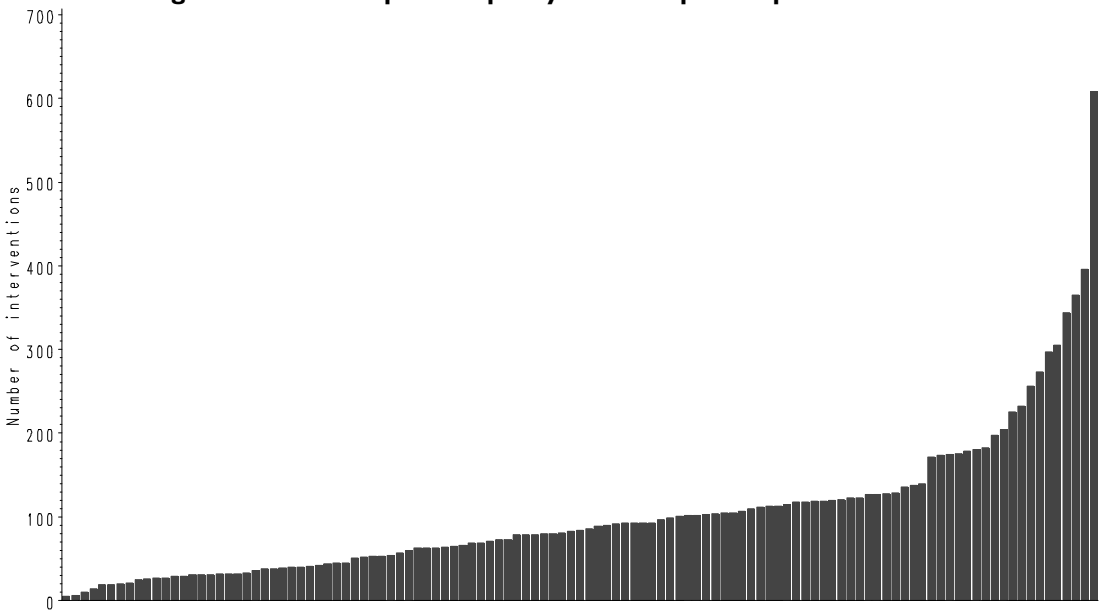


Table 7.7 gives information on the number of elective THR per surgeon. For surgeons who work in several hospitals, a sum was made of all their THR. On average, a surgeon performed 23 elective THR in 2004. The maximum was 250 elective THR. Three quarters of surgeons performed 27 elective THR or less.

Table 7.7: Total hip arthroplasty: Summary measures of volume per surgeon

Number	Minimum	Mean	25th Pctl	50 Pctl	75th Pctl	Maximum
522	1.0	22.7	4.0	11.0	27.0	250.0

Figure 7.2: Total hip arthroplasty: Volume per surgeon

7.1.5 Definition of outcomes

Common outcome measures are mortality, loss of independence, loss of mobility and residual pain. Unfortunately, the latter three are not registered in the MCD. Complications associated with THR can be divided in two kinds. Complications associated with immobilization include development of deep vein thrombosis, pulmonary embolism, pneumonia, and muscular atrophy. Complications associated with the surgical procedure include postoperative infection, mechanical malfunctions of the prosthesis and dislocation of the hip. These complications can give cause for a re-intervention, i.e. a revision hip replacement. Finally, in the long term, there is the problem of hip replacements progressively wearing out over the years. Patients who wear out their prosthesis will require a revision hip replacement surgery.

The literature review (see Table 2.1 on page 16) identified the following outcome measures for THR: in-hospital mortality, in-hospital complication rate, 90-day complication rate and 90-day revision rate. The revision rate was also studied at 48 months, but since MCD data were only available for 2004 and 2005, this long-term outcome measure could not be analysed in this report.[62]

7.1.5.1 In-hospital outcome

In-hospital outcome includes in-hospital mortality as well as complications after total hip arthroplasty. Information on in-hospital complications was retrieved from the MCD, on the basis of the following ICD-9-CM secondary diagnosis codes during index admission:

- Pulmonary embolism: 415.1 "Pulmonary embolism and infarction"
- Deep venous thrombosis (DVT): 451.1x "Phlebitis and thrombophlebitis of deep vessels of lower extremities"
- Deep wound infection (SSI for surgical site infection): 996.66 "Infection and inflammatory reaction due to internal joint prosthesis". (Code 998.5 "Postoperative infection" was not retained as complication in the present study, as this adverse event was supposed to be linked to the general quality of care in the hospital and not to a specific procedure.)

These outcome definitions were derived from the KCE report on quality indicators by Vlayen and colleagues.[310]

Those four in-hospital outcomes are very rare events, as shown in Table 7.8. Therefore, no further analysis in terms of VOA was performed on in-hospital outcomes and all our attention was focused on events occurring within 90 days after the first index stay, and on the revision rate at 18 months.

Table 7.8: Total hip arthroplasty: In-hospital outcomes (index stay)

	Number total	number outcome	% outcome
Death	11856	28	0.24
Deep venous thrombosis	11856	12	0.10
Pulmonary embolism	11856	12	0.10
Deep wound infection	11856	7	0.06

7.1.5.2 Complications within 90 days

The 90 days interval starts with the total hip replacement index admission. This period of time was chosen previously by Katz. et al. to maximise the likelihood of causality between the hip replacement and the adverse event as effect of the procedure.[311]

Information on the 90-day complication rate is retrieved from two sources:

1. the index admission: on the basis of the secondary diagnosis;
2. consecutive readmissions that occurred within 90 days from index admission: on the basis of the principal diagnosis or procedure.

The following ICD-9-CM diagnosis (as primary or secondary diagnosis) or procedure codes are used:

- Pulmonary embolism: 415.1 "Pulmonary embolism and infarction as principal diagnosis"
- Deep venous thrombosis (DVT): 451.1x "Phlebitis and thrombophlebitis of deep vessels of lower extremities"
- Deep wound infection: 996.66 "Infection and inflammatory reaction due to internal joint prosthesis"
- Mechanical complication: 996.4 "Mechanical complication of internal orthopaedic device, implant, & graft"
- Other complications: at least one ICD-9-CM diagnosis code indicating "Other complications of internal prosthetic device, implant, & graft":
  - 996.77 "Due to internal joint prosthesis"
  - or 996.78 "Due to other internal orthopaedic device, implant, & graft"
  - or 996.79 "Due to other internal prosthetic device, implant, & graft"
- Dislocation of hip within 90 days from initial admission: at least one of the following codes:
  - diagnosis 835.x "Dislocation of hip"
  - or procedure 79.75 "Closed reduction of dislocation of hip"
  - or procedure 79.85 "Open reduction of dislocation of hip".

As shown in Table 7.9, the most frequent complication observed during hospitalization (initial or readmissions) within 90 days after admission was a mechanical complication of the prosthesis with 1.96% 95%CI (1.72%, 2.22%), followed by 1.69% of hip dislocation 95%CI (1.46%, 1.94%). There is a certain overlap between categories. For example, some patients who were readmitted in a hospital with a principal diagnosis 996.4 "Mechanical complication" also underwent a reduction of hip dislocation (ICD-9-CM procedure 79.75 or 79.85). Some patients also experienced different complications or were readmitted more than once. Therefore, a global outcome category was created to group all patients who encountered any of the above complications within 90 days after the initial hip arthroplasty admission. Table 7.9 shows that 3.33% 95%CI (3.02% 3.67%) of the patients (number = 395) had at least one of the studied post-operative complications within 90 days after the first THR.

Table 7.9: Total hip arthroplasty: Outcomes measured during hospital (re)admission within 90 days of index admission

Complication within 90 days	Number total	number outcome	% outcome	95% CI	
Deep venous thrombosis	11856	16	0.13	0.08	0.22
Pulmonary embolism	11856	45	0.38	0.28	0.51
Deep wound infection	11856	21	0.18	0.11	0.27
Mechanical complication	11856	232	1.96	1.72	2.22
Dislocation of the hip	11856	200	1.69	1.46	1.94
Other complication	11856	56	0.47	0.36	0.61
Any of the above complications (*)	11856	395	3.33	3.02	3.67

(*) a patient may have more than one complication, but these are counted only once

7.1.5.3 Revision rate at 18 months

Revision rates at 18 months were retrieved by identifying the readmissions presenting an ICD-9-CM procedure code 8153 OR a NIHDI procedure code 293440 (see Table 5.2).

Figure 7.3 shows the Kaplan-Meier survival curve: the vertical axis gives the proportion of people free from revision; the horizontal axis gives the time (in days) after the first THR. As shown in Table 7.10, the cumulative percentage of revision at 18 months (converted into 547 days) was estimated at 1.84% 95% CI (1.6%, 2.11%) using a Kaplan-Meier survival curve.

Revisions at 48 months (4 years), as presented in the literature, were not calculated since data were not available.

Figure 7.3: Total hip arthroplasty: Patients free from revision (survival function, Kaplan-Meier)

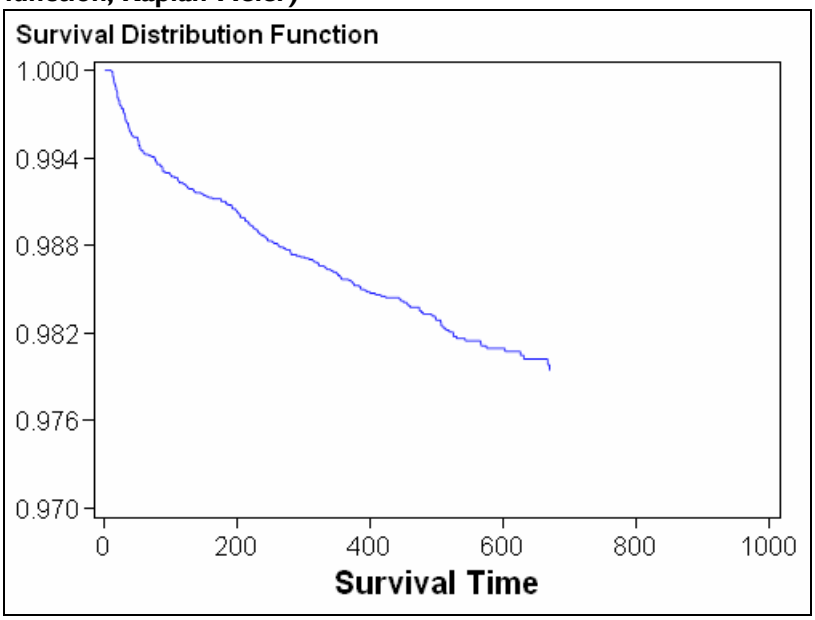


Table 7.10: Total hip arthroplasty: Revision rate at 18 months

	Number total	Number events	Number censored	% censored	Kaplan-Meier estimator	95% CI	
Readmission for revision at 18 months	11856	214	11642	98.20%	1.84%	1.60%	2.11%

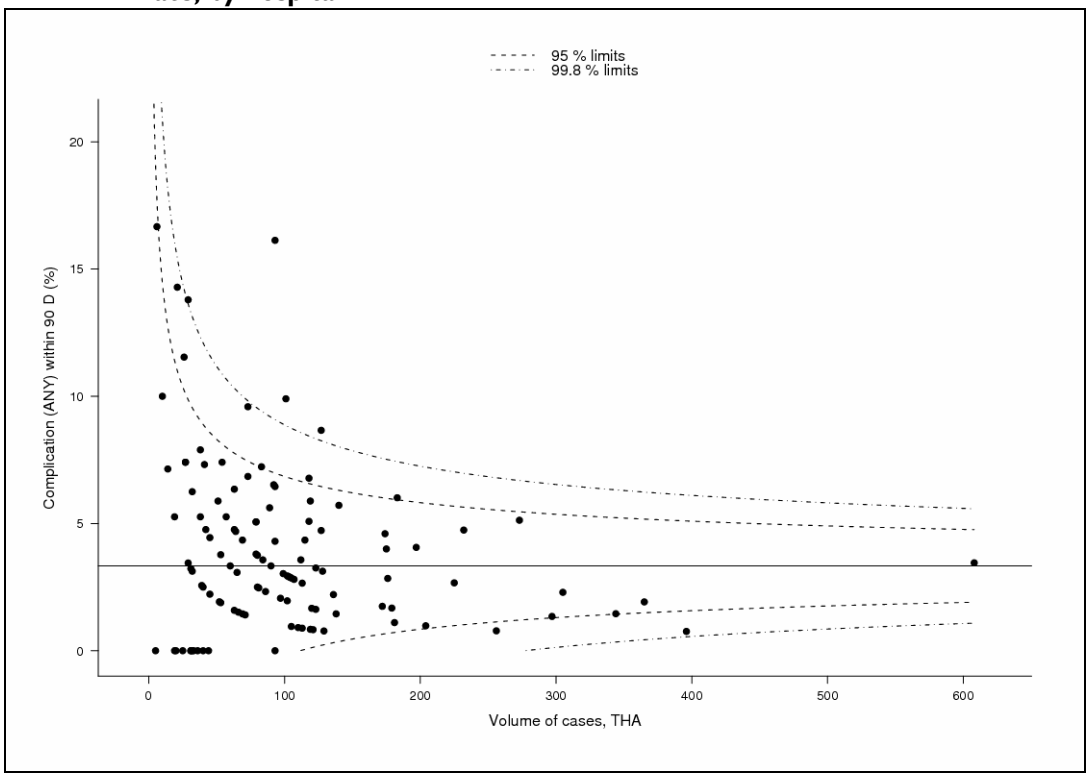
7.1.6 Volume outcome relationship

7.1.6.1 Analysis by hospital

90-DAY COMPLICATION (ANY) RATE, BY HOSPITAL

The funnel plot of the relationship between complication rate within 90 days and volume of interventions is presented in Figure 7.4. The horizontal line represents the overall 90-day complication rate i.e. 3.3% (for all hospitals). A few small hospitals had 0% complications. Three hospitals with less than 200 procedures may be considered as outliers above the 99.8% limits of variability. Among those 3 hospitals, one hospital had a high percentage of dislocation, either during the index stay (6/93) or during the 90 days after admission (13/93) that explains the high rate (15/93 = 16%) of complication within 90 days. The hospital with the highest volume was located on the national rate, inside the limits.

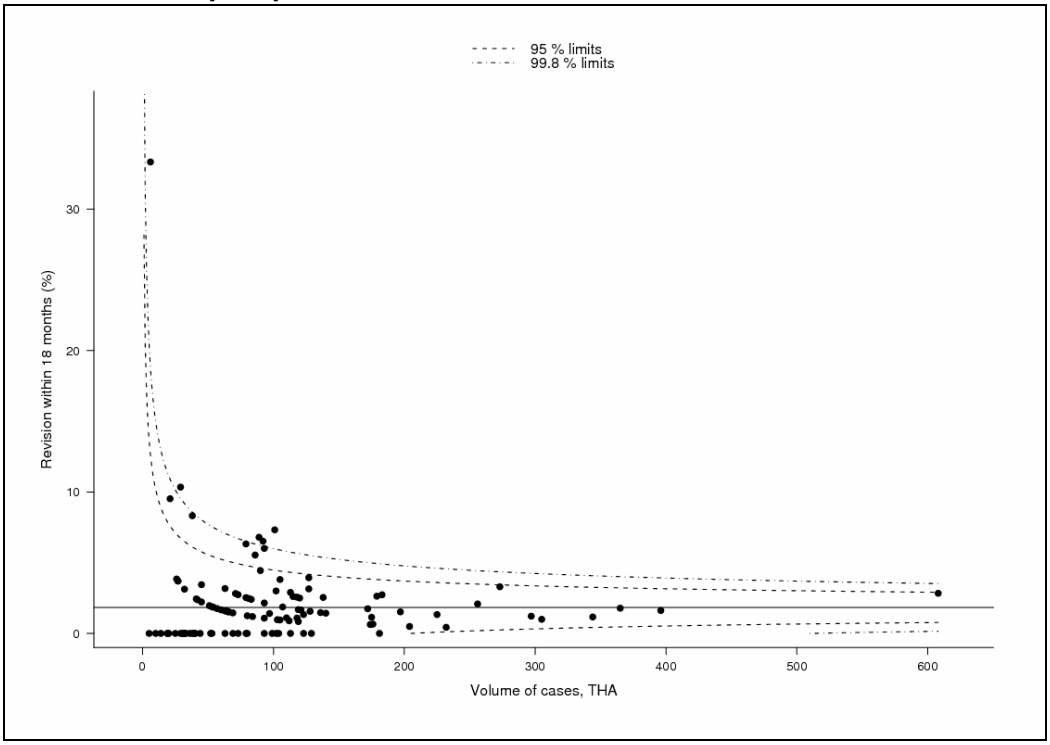
Figure 7.4: Total hip arthroplasty: Funnel plot of the 90-day complication rate, by hospital



REVISION RATE AT 18 MONTHS, BY HOSPITAL

The funnel plot of the relationship between revision rate within 18 months and volume of interventions is presented in Figure 7.5. A few small hospitals were located above the 99.8% boundaries.

Figure 7.5: Total hip arthroplasty: Funnel plot of the revision rate at 18 months, by hospital



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

Table 7.11 presents the differences in case-mix in function of the volume of total hip arthroplasty performed per hospital. Low-volume hospitals (in which less than 61 THR were performed in 2004) have slightly more male and older patients, and the latter have more co-morbidities (i.e. they have a higher Charlson index score). This was confirmed by the fact that, for principal diagnosis 715, the severity of illness seemed to be higher in these small hospitals. The ICD-9-CM principal diagnosis 715 "Osteoarthrosis" was less frequently coded in the small hospitals, in favour of code 733 "Other disorders of bone and cartilage". This could reveal a different case-mix or be due to a less specific coding behaviour in small hospitals.

Table 7.11: Total hip arthroplasty: Differences in case-mix and outcomes by volume category of hospital

		Tertile based on volume			
		≤60 THR	> 60 and ≤110 THR	> 110 THR	All
Number hospitals		39	38	38	115
Number stays		1327	3248	7281	11856
Male	%	42.0	38.0	39.5	39.4
Age	Mean	69.0	68.2	67.1	67.6
	Std	11.1	11.3	11.8	11.6
Charlson index score	Mean	0.6	0.4	0.4	0.4
	Std	1.1	1.0	0.9	1.0
Charlson ≥ 3	%	6.8	4.2	3.3	3.9
Principal diagnosis					
715 Osteoarthrosis	%	86.7	92.3	92.4	91.7
733 Other disorders of bone & cartilage	%	8.4	6.3	6.3	6.6
Other principal diagnosis	%	4.8	1.4	1.3	1.7
Severity index (APR-DRG) for stays with principal diagnosis 715		1151	2997	6725	10873
1	%	42.9	51.1	52.9	51.3
2	%	42.7	40.8	41.2	41.2
3	%	13.5	7.3	5.3	6.7
4	%	0.9	0.8	0.7	0.7
Complication rate (any) within 90 days	%	4.1	4.1	2.9	3.3
Revision rate at 18 months	%	1.9	2.3	1.7	1.8

RELATION BETWEEN HOSPITAL VOLUME AND 90-DAY COMPLICATION (ANY) RATE

The effects from the logistic regression, with the 90-day complication rate as dependent variable, are presented in Table 7.12. Without taking case-mix into account, an increase of volume of 10% is associated with a statistically significant change in the odds of a complication of -2.62% 95% CI (-4.30%, -0.94%). After adjustment for the case-mix, the change is reduced to -1.63% 95% CI (-3.43%, 0.17%) and is not statistically significant anymore. On the basis of these results, we therefore conclude that, taking case-mix into account, there is no association between the hospital volume of THR and the 90-day complication rate.

Table 7.12: Total hip arthroplasty: Correlation-corrected logistic regression (GEE) estimates of determinants of 90-day complication

Model without adjustment for case mix	**Effect[1]**	**95% CI**	
Hospital volume (increase of 10%)	-2.62	-4.30	-0.94
Model with adjustment for case mix	**Effect[1]**	**95% CI**	
Hospital volume (increase of 10%)	-1.63	-3.43	0.17
	Odds Ratio	95% CI	
Sex (male vs. female)	1.05	0.84	1.32
Age (increase of 1 year)	1.02	1.01	1.03
Charlson score (increase of 1 category)	1.48	1.30	1.69
715 vs. other PDX	0.70	0.35	1.39
733 vs. other PDX	1.14	0.57	2.29

[1] Effect of 10% increase in volume of hospital on the odds of complication within 90 days

RELATION BETWEEN HOSPITAL VOLUME AND REVISION RATE AT 18 MONTHS

Table 7.13 presents the results of the Cox proportional hazards model. When taking the case-mix into account (i.e. adjusting for age, gender, Charlson score and principal diagnosis), medium-volume hospitals have a statistically significant higher revision rate at 18 months than high-volume hospitals: 1.451 95% CI (1.004, 2.098). There was, however, no effect when small and high volume hospitals were compared: 1.137 95% CI (0.701, 1.846). These opposite results are difficult to understand.

Table 7.13: Total hip arthroplasty: Results of regression: hazard ratio of hospital volume category on revision rate at 18 months

Model without adjustment for case mix	Effect[1]	95% CI	
Volume: low-volume (<60 THR) versus high-volume hospital (>110 THR)	1.103	0.678	1.794
Medium-volume versus high–volume hospital	1.427	0.986	2.066
Model with adjustment for case mix	Effect[1]	95% CI	
Volume: low-volume (<60 THR) versus high-volume hospital (>110 THR)	1.137	0.701	1.846
Medium-volume versus high–volume hospital	1.451	1.004	2.098
	Hazard Ratio	95% CI	
Sex (male vs. female)	1.146	0.877	1.498
Age (increase of 1 year)	0.988	0.977	0.998
Charlson score (increase of 1 category)	0.969	0.768	1.222
715 vs. other PDX	1.992	0.506	7.841
733 vs. other PDX	2.791	0.650	11.976

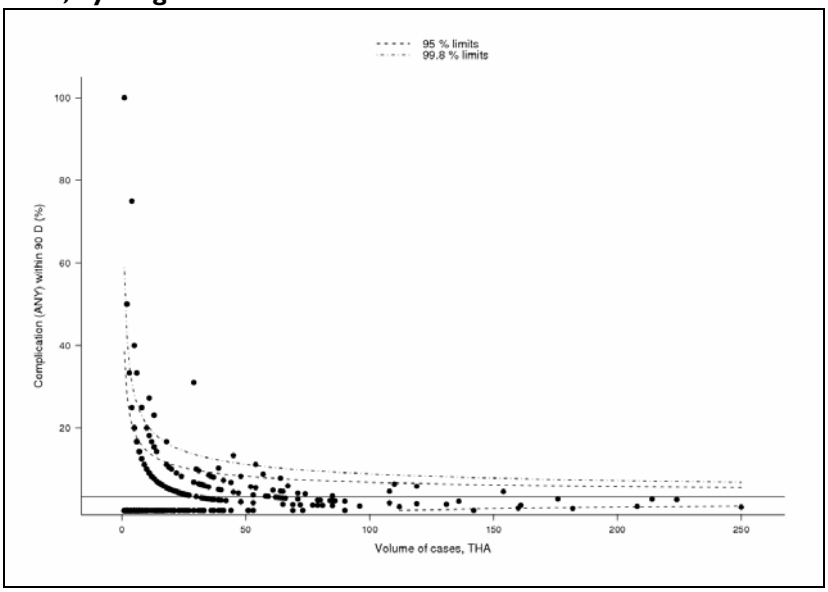
[1] Hazard ratio of volume category on revision within 18 months

7.1.6.2 Analysis by surgeon

90-DAY COMPLICATION (ANY) RATE, BY SURGEON

The funnel plot (Figure 7.6) of the relationship between the complication rate within 90 days and the volume of interventions is presented in Figure 6.2. About 9 surgeons with less than 50 elective THR were located outside the 99.8% limits and may therefore be considered as outliers.

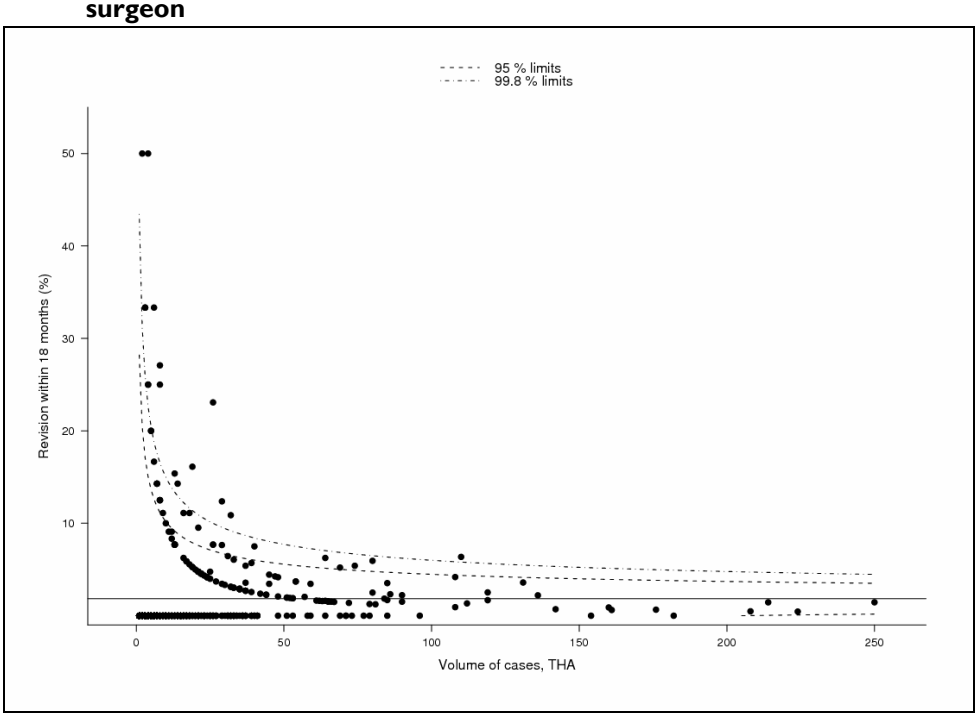
Figure 7.6: Total hip arthroplasty: Funnel plot of the 90-day complication rate, by surgeon



REVISION RATE AT 18 MONTHS, BY SURGEON

As presented on the funnel plot in Figure 7.7, about 12 surgeons were located outside the 99.8% limits of the distribution of the revision rate at 18 months; all but one had a volume of less than 50 elective THR in 2004.

Figure 7.7: Total hip arthroplasty: Funnel plot of the revision rate at 18, by surgeon



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE SURGEON

Table 7.14 shows that surgeons who performed less than 6 elective THR in 2004 have higher risk patients; they have a higher percentage of patients with a Charlson score of three or more, and more patients with a severity of illness of three. Complication rate within 90 days and revisions at 18 months was 5.0% for small volume surgeons and 3.0% for high volume surgeons.

Table 7.14: Total hip arthroplasty: Differences in case mix and outcomes by volume category of surgeon

		Tertile based on volume			
		≤6 THR	6 THR > and ≤20 THR	>20 THR	All
Number surgeons		190	159	173	522
Number stays		604	1960	9292	11856
Male	%	40.6	38.0	39.6	39.4
Age (years)	Mean	67.4	68.5	67.4	67.6
	Std	12.9	11.4	11.5	11.6
Charlson index score	Mean	0.5	0.5	0.4	0.4
	Std	1.3	1.0	1.0	1.0
Charlson ≥ 3	%	6.0	4.9	3.6	3.9
Principal diagnosis					
715 Osteoarthrosis	%	82.5	89.9	92.7	91.7
733 Other disorders of bone & cartilage	%	11.8	8.5	5.8	6.6
Other principal diagnosis	%	5.8	1.5	1.5	1.7
Severity index (APR-DRG) for stays with principal diagnosis 715		498	1763	8612	10873
1	%	47.4	47.0	52.5	51.3
2	%	40.0	43.5	40.9	41.2
3	%	10.6	8.6	6.1	6.7
4	%	2.0	0.9	0.6	0.7
Complication rate (any) within 90 days	%	5.0	4.5	3.0	3.3
Revision rate at 18 months	%	2.01	2.04	1.80	1.84

RELATION BETWEEN SURGEON VOLUME AND 90-DAY COMPLICATION (ANY) RATE

Results of the logistic regression are given in Table 7.15. Without taking case-mix into account, an increase of volume operated by the surgeon of 10% is associated with a statistically significant change in the odds of complication of -3.13% (-4.34%, -1.92%). After adjustment for the case-mix, the change stays statistically significant but is reduced to -2.68% (-3.92%, -1.45%). Actually, the Charlson score is the only other variable to have a statistically significant effect on the outcome.

Table 7.15: Total hip arthroplasty: Correlation-corrected logistic regression (GEE) estimates of determinants of 90-day complication

Model without adjustment for case mix	Effect[1]	95% CI	
Surgeon volume (increase of 10%)	-3.13	-4.34	-1.92
Model with adjustment for case mix	Effect[1]	95% CI	
Surgeon volume (increase of 10%)	-2.68	-3.92	-1.45
	Odds Ratio	95% CI	
Sex (male vs. female)	1.08	0.87	1.34
Age (increase of 1 year)	1.01	1.00	1.03
Charlson score (increase of 1 category)	1.43	1.24	1.65
715 vs. other PDX	0.68	0.35	1.34
733 vs. other PDX	1.09	0.53	2.23

[1] Effect of 10% increase in volume of surgeon on the odds of complication within 90 days

RELATION BETWEEN SURGEON VOLUME AND REVISION RATE AT 18 MONTHS

The effects from the Cox proportional hazards model, with revision rate at 18 months as dependent variable, are presented in Table 7.16. Low volume surgeons and medium volume surgeons have a slightly higher revision rate when compared to high volume surgeons: hazard ratio and 95%CI 1.213 (0.641, 2.295) for low volume compared to high volume and 1.156 (0.782, 1.709) for medium volume compared to high volume, respectively.

Table 7.16: Total hip arthroplasty: Results of regression: hazard ratio of surgeon volume category on revision rate at 18 months

Model without adjustment for case mix	Effect[1]	95% CI	
Volume: low-volume (<6 THR) versus high-volume surgeon (>20 THR)	1.221	0.648	2.302
Medium-volume versus high–volume surgeon	1.148	0.779	1.693
Model with adjustment for case mix	Effect[1]	95% CI	
Volume: low-volume (<6 THR) versus high-volume surgeon (>20 THR)	1.213	0.641	2.295
Medium-volume versus high–volume surgeon	1.156	0.782	1.709
	Hazard Ratio	95% CI	
Sex (male vs. female)	1.145	0.865	1.516
Age (increase of 1 year)	0.988	0.976	1.001
Charlson score (increase of 1 category)	0.968	0.742	1.264
715 vs. other PDX	2.034	0.507	8.163
733 vs. other PDX	2.814	0.667	11.868

[1] Hazard ratio of volume category of surgeon on revision within 18 months

7.1.7 Discussion

7.1.7.1 External validation of the definition of the procedure

The external validation of our definition process is based on the data from the Alliance of Christian Sickness Funds in which a total of 17 485 total hip replacements were reported in Belgium in 1998.[312, 313] This number was obtained by extrapolation of the claims data from the Christian Sickness Funds for the whole of Belgium. It included elective and emergency THR and should therefore be compared with the 20 267 stays selected on the basis of NIHDI procedure codes (see Table 5.2 on page 46). The Christian Sickness Funds mentioned a yearly increase of 4.5% of the number of elective THR and of 2% of the emergency THR, in Belgium, in the period 1990-1998. When such a yearly increase is taken into account, their data become very similar to ours.

7.1.7.2 Summarized results of literature review

The systematic literature search identified six systematic reviews in which the VOA for THR was studied.[1, 5, 59, 60, 62, 64] In all, these systematic reviews identified 20 primary studies of which 4 were published between 2001 and 2004, [311, 314-316] and 7 between 1995 and 1999.[197, 309, 317-321] These articles were complemented with a limited number of primary studies that were retrieved through the reference lists of the original 20 primary studies and through a search of grey literature.[312, 322, 323]

On the basis of all six systematic reviews, it was concluded in Chapter 2 (see Table 2.2 on page 19) that there were conflicting results in relation to the volume-outcome association between hospital volume and in-hospital mortality. The term "conflicting results" is used because there is a mix of primary studies that indicate an inverse relation with volume, others that indicate no relation, and others that indicate a direct relation with volume (which means that an increase in volume is associated with higher mortality). It was emphasised, however, that each systematic review had studied only a very limited number of primary studies.

Since there was only one systematic review which studied other outcome measures such as revision rate and 90-day complication rate, these results could not be summarized in the evidence table in Appendix 10.[62]

Detailed results of the literature review, using the data extraction template, are described in the evidence tables in Appendix 10.

7.1.7.3 Comparative analysis of literature and Belgian data

DEFINITION OF VOLUME

As shown in Table 7.6, Belgian hospitals had a mean annual elective THR volume of 103 in 2004; median is 84; 75th Percentile is 121. These volumes seem higher than in the US, where 44% of US hospitals performed ten or fewer elective THR, in 2003.[323]

Table 7.7 shows that Belgian orthopaedic surgeons performed on average 23 elective THR in 2004; median is 11; 75th Percentile is 27 THR per year. These annual volumes per surgeon seem rather high when compared with foreign data. Zahn et al. calculated that 28% of all US elective THR in 2003 were performed by surgeons who carried out five or fewer of these procedures annually.[323] In Belgium, in 2004, only 5% of elective THR were performed by surgeons with six or fewer elective THR (see Table 7.14). It has to be mentioned that elective surgery can be supply induced.

OUTCOME

A great disadvantage of the MCD is the fact that they do not capture clinically relevant outcome measures such as loss of dependence, loss of mobility or residual pain.

In addition, it has to be acknowledged that the Charlson score is an inappropriate score for risk adjustment. For orthopaedic surgery, the Physiological and Operative Severity Score for enUmeration of Mortality and morbidity (POSSUM) adapted for orthopaedic patients was assessed as a useful tool to predict morbidity and mortality.[324] However, the orthopaedic POSSUM system includes a physiological assessment and an analysis of operative severity; information that is not available in the MCD.

In-hospital mortality after total hip replacement

Although it was decided not to study the volume-outcome association for in-hospital mortality after THR, it remains interesting to compare this study's unadjusted percentages with other studies. The in-hospital mortality rate associated with THR was 0.24% in this KCE study (see Table 7.8). Zahn et al. used 2003 nationwide United States data and they reported a similar in-hospital mortality rate of 0.33% (95% CI 0.27%, 0.36%).[323]

90-day complication rate after total hip replacement

Table 7.17 gives an overview of the rates that were found in the primary studies, in comparison with those found in this KCE study. Overall, the 90-day complication rates after THR are lower in Belgian administrative data than in those published by Katz and Solomon. Katz et al. published 90-day postoperative outcomes based on Medicare US claims data for 1995-1996.[311] The complication rates were respectively 3.13% (hip dislocation), 0.23% (deep wound infection) and 0.93% (pulmonary embolism). Solomon et al. used a subset of Katz's data and found 0.4% postoperative deep wound infection within 90 days and 2.3% hip dislocation within 90 days.[315]

Table 7.17: Total hip arthroplasty: 90-day complication rate: comparison of Volume-Outcome KCE study with scientific literature

COMPLICATION WITHIN 90 DAYS	KCE study Volume outcome	Katz 2001 USA	Solomon 2002 USA
Pulmonary embolism	0.38%	0.93%	-
Deep wound infection	0.18%	0.23%	0.4%
Dislocation of hip	1.69%	3.13%	2.3%

There are several possible explanations for these differences.

First, deep wound infection was defined differently by Katz and Solomon. They only included deep wound infections that required either surgical debridement or removal of the prosthesis, while the KCE study searched for the diagnostic code 996.66 "Infection and inflammatory reaction due to internal joint prosthesis".

Second, the low 90-day complication rates are probably related to the fact that complications are not well coded in the MCD. As it happens, a previous KCE study assessed the accuracy of Belgian MCD for detecting adverse events. Gillet et al. [111] concluded that deep venous thrombosis or pulmonary embolism and postoperative wound infection were no valid indicators for an adequate detection of these adverse events. In case of pulmonary embolism and deep venous thrombosis, these results are surprising since the diagnostic codes 415.1 and 451.1x have an initial level of severity of, respectively, 3 and 4. Omitting to code them is not in the interest of the hospital since the Belgian hospital financing system is based on the hospital case-mix. The diagnostic code for deep wound infection (996.66), on the other hand, has an initial level of severity that equals 1, which means that coding it cannot modify the stays global level of severity. On the basis of these practical implications, one would expect that hospitals would be financially motivated to encode DVT and pulmonary embolism. The report by Gillet et al. proved this assumption to be wrong.[111]

Third, Katz and Solomon used both inpatient as outpatient claims data.

Fourth, Katz and Solomon took into account secondary diagnoses coded during readmissions to calculate the complication rate, whereas the KCE study did not include the secondary diagnoses. As the population might have needed orthopaedic care for other joints than the operated hip due to their degenerative disease or arthritis, we chose not to select complication codes present as secondary diagnoses in posterior stays. Consequently, the reliability of the link to the hip placement was preferred to the detriment of a possible underestimation of the rate.

Fifth, the selection process applied in our study focused on single total hip arthroplasty, excluding bilateral procedures and patients with more than one procedure during 2004-2005. These patients stayed included in the US studies.

Finally, Katz and Solomon, on the other hand, excluded patients who were less than 65 years old.

Revision rate at 18 months after total hip replacement

Table 7.18 shows similar revision rates at 18 months in different studies. The 1.84% revision rate in the 2004 Belgian MCD (see Table 7.10) is very comparable to the +/- 2% which was observed by Losina et al. in the Medicare claims data for 1995-1996.[314] In 2000, Diels et al. calculated an 18 month revision rate of approximately 1.5% among all members of the Belgian Christian Sickness Funds who underwent a total hip replacement between 1990 and 1999.[312]

Table 7.18: Total hip arthroplasty: revision rate at 18 months: comparison of Volume-Outcome KCE study with other studies

COMPLICATION WITHIN 18 MONTHS	KCE study Volume outcome	Losina 2004 USA	Diels 2000 Belgium
Revision rate at 18 months	1.84%	+/- 2%	+/- 1.5%

PATIENT CASE MIX

Table 7.11 and Table 7.14 showed that low-volume hospitals (≤60 THR/year) and low-volume surgeons (≤6 THR/year) have higher risk patients based on a higher Charlson score and a higher severity of illness. Katz and colleagues made similar observations in the course of their population-based cohort study of patients who underwent elective THR in 1995.[322] They found that patients who underwent surgery in higher-volume hospitals (>100 elective THR/year) were younger, and were significantly more likely to have better recalled preoperative functional status. In addition, these patients were more likely to have attended college and to have an income greater than $20 000 per year.

VOLUME OUTCOME ASSOCIATION

RELATION BETWEEN HOSPITAL OR SURGEON VOLUME AND 90-DAY COMPLICATION (ANY) RATE

After adjustment for age, gender, Charlson score and principal diagnosis, higher hospital volume was associated with a lower rate of complications within 90 days after the index THR, but the trend did not reach significance (see Table 7.12). Higher surgeon volume, on the other hand, was statistically significant associated with a lower 90-day complication rate (see Table 7.15).

Katz and colleagues assessed the relationship between surgeon and hospital procedure volume and mortality, hip dislocation, deep wound infection (requiring surgical debridement or removal of the prosthesis) and pulmonary embolism in the first ninety days postoperatively.[311] This assessment was done with claims data of Medicare patients who underwent an elective THR in 1995-1996. Analyses were adjusted for age, gender, arthritis diagnosis, a proxy for low income and the Charlson co-morbidity index. In addition, analyses of hospital volume were adjusted for surgeon volume, and analyses of surgeon volume were adjusted for hospital volume. Higher hospital volume was significantly associated with lower rates of 90-day mortality and hip dislocation after primary total hip replacement (p value <0.01 for each outcome). Higher surgeon volume was significantly associated with a lower rate of hip dislocation (p value = 0.0001). The effect of surgeon volume was strongest in hospitals with ≤100 THR/year. In hospitals with more than 100 cases annually, surgeon volume had little effect on outcomes.

Solomon and al. used a subset of Katz's data and administered a survey to hospital administrators to obtain information on hospital characteristics.[315] Their objective was to examine whether specific hospital-level characteristics could account for the association between the volume of THR and the 90-day rate of hip dislocation and deep wound infection of the hip. They found that only the volume of THR performed by an individual surgeon was an independent predictor of adverse events and that hospital volume and hospital-level factors had no appreciable association with adverse events. Patients undergoing THR in low-volume hospitals whose surgeons performed ≥10 THR/year had a lower risk of adverse events compared with patients whose surgeons performed <10 THR per year in the same hospitals. Likewise, patients whose surgeons performed <10 THR per year in high-volume hospitals had a higher risk of an adverse event than patients in the same hospitals whose surgeons performed ≥10 THR/year.

Although there are methodological differences between our KCE study and the primary studies by Katz and Solomon (see page 182), they seem to reach similar conclusions on the inverse association between surgeon volume and complication rate after THR.

RELATION BETWEEN HOSPITAL OR SURGEON VOLUME AND REVISION RATE AT 18 MONTHS

After adjustment for age, gender, Charlson score and principal diagnosis, medium-volume hospitals (between 60 and 110 THR per year) have a statistically significant higher revision–risk at 18 months than high-volume hospitals (>110 THR/year). However, there is no effect when small and high volume hospitals were compared. See Table 7.13. These results are hard to understand. For surgeons, there is no relation between revision rate and surgeon volume (see Table 7.16).

Losina and colleagues found that patients operated upon by low-volume surgeons were considerably more likely to undergo a revision of the THR than patients operated upon by high-volume surgeons. This association occurred primarily during the first 18 months after the index THR. It is difficult, however, to compare Losina's results with ours because of the use of other volume thresholds for surgeons and hospitals. Low-volume surgeons were defined as those performing <12 THR/year; high-volume surgeons as those performing ≥12 THR/year. For high-volume surgeons, they found an association between the risk of THR failure and hospital volume: hospitals with an annual caseload of >25 THR/year had lower failure rates than hospitals with less THR. For low-volume surgeons, there was no association between revision rate and hospital volume.[314]

Espehaug and al. found a similar association between hospital volume and revision rate at 4 years (threshold at 11 THR/year). In addition, they investigated the effect of cemented versus uncemented prostheses and found that the volume effect was primarily seen in patients who received uncemented prostheses.[318]

Diels and colleagues also found an association between hospital volume and revision rate; when the hospital volume increased with one THR procedure, the revision risk decreased with 0.3%. Their analyses also showed that patients with an uncemented prosthesis had 50% more risk of a revision than patients with a cemented monobloc inox-prosthesis.[312]

RELATION BETWEEN HOSPITAL OR SURGEON VOLUME AND POSTOPERATIVE FUNCTIONAL STATUS AND PAIN RELIEF

Although these outcome measures could not be studied in the Belgian MCD, it seems important to discuss the results of other studies on this topic to provide a more complete picture of the influence of volume on outcome.

Katz and colleagues evaluated whether hospital and surgeon volume of THR are associated with patient-reported pain and functional status (the so-called Harris hip score) and satisfaction with surgery 3 year postoperatively.[322] As mentioned earlier, Katz et al. observed that patients with low levels of income and education and those with worse recalled preoperative functional status, were more likely to have THR performed at low-volume hospitals. Before adjusting for these factors, low hospital volume was associated with worse Harris hip scores at follow-up. After adjustment for socio-demographic and clinical variables, however, the association between higher hospital volume and better functional status following primary THR was weak and statistically non-significant. Satisfaction with primary THR, on the other hand, remained greater among patients whose operations were performed in higher-volume hospitals.

Thompson and colleagues used medical records and questionnaires to assess the association between hospital and surgeon volume of elective THR and several outcomes i.e. in-hospital operative complications, in-hospital general complications, 6-month difficulty walking and 6-month residual pain.[316] Contrary to most other studies that are limited to claims data, Thompson disposed of a variety of preoperative clinical risk factors for case-mix adjustment, i.e. activity level, ADL scale, ASA score, walking distance and hip pain score. He also made separate analyzes for patients receiving cementless and cemented prostheses. Thompson et al. did not conclude that, in general, surgeon nor hospital volume had any significant association with the likelihood of operative or general complications, nor with walking and pain outcomes.

Key points on volume outcome association for elective total hip replacement/arthroplasty (THR)

- **A total of 11 856 patients were hospitalized in 2004 for a total hip arthroplasty. Partial arthroplasties, revisions of hip replacement, bilateral procedures and procedures for trauma were not included in this selection.**
- **These interventions were performed in 115 centres by 522 surgeons. Annual mean number of THR was approximately 100 per centre, and 23 per surgeon. These volumes are much higher than US data (44% of centres less than 11 elective THR).**
- **In-hospital mortality is virtually inexistent after elective THR, and in-hospital complications are also very rare (or under reported in the MCD). Following outcomes were analysed: complications requiring a hospitalisation within 90 days of intervention (3.3% of all patients), and revision of the prosthesis at 18 months (1.8% of all patients). Logistic and Cox regression models were fitted to assess the association between these outcomes and hospital or surgeon volume. Factors included in the models were: age, sex, Charlson score and principal diagnosis (ostheoarthrosis, other disorders of bone or cartilage, or other).**
- **The literature review did not succeed in identifying good quality systematic reviews that studied revision rate and 90-day complication rate after THR.**
- **Complication rate at 90 days was 4.1% for hospitals with less than 110 THR/year, and 2.9 % for other centres. Results from logistic regression without adjustment for case mix showed a statistically significant effect of hospital volume on the probability of complication rate (increase of 10% of volume associated with decrease of 2.6% in odds of death). After adjustment for case mix, this effect was reduced (1.6% decrease) and was only borderline statistically significant.**
- **The association between surgeon volume and likelihood of complication at 90 days was more robust: 5.0% for surgeons with ≤ 6 THR/year, 4.5% for surgeons between 6 and 20 THR/yr, and 3.0% for surgeons with more than 20 THR/year. This association remained statistically significant after adjustment for case mix.**
- **Revision rates at 18 months were lowest in high volume centres (1.7% in centres with more than 110 THR/year), highest (2.9%) in medium volume centres (between 60 and 110 THR) and 1.9% in low volume centres (less than 60 THR/year). Increases in volume were thus not consistently associated with decreases in revision rate.**
- **A small decrease was observed in revision rates at 18 months from small or medium volume surgeons (≤ 20 THR/year, 2.0%) in comparison with high volume surgeons (> 20 THR/year, 1.8%), but effects are too small to draw conclusions.**
- **Unfortunately, the MCD do not provide information on the outcomes of greatest interest to patients such as loss of independence, loss of mobility or residual pain.**
- **POSSUM (Physiological and Operative Severity Score for enUmeration of Mortality and morbidity) would be better for risk adjustment in orthopaedic surgery than the applied Charlson score. However, these clinical parameters are not encoded in the MCD.**

7.2 ELECTIVE TOTAL KNEE REPLACEMENT/ARTHROPLASTY (TKR)

7.2.1 General description of procedure

Total knee replacement or arthroplasty is a procedure in which damaged or diseased parts of the knee joint are removed and replaced with an artificial joint, called prosthesis. The thigh bone (or femur) abuts the large bone of the lower leg (tibia) at the knee joint. During a total knee replacement, the end of the thigh bone (femur) is removed and replaced with a metal shell. The end of the lower leg bone (tibia) is also removed and replaced with a channelled plastic piece with a metal stem. Depending on the condition of the kneecap, a plastic "button" may also be added under the kneecap surface.

Revision knee replacement is performed when the original primary total knee replacement has worn out or loosened in the bone, or when the primary TKR fails due to, for example, recurrent dislocation, infection or fracture.

This report only studies the volume-outcome relationship for the primary total knee replacement.

7.2.2 Primary data selection in MCD 2004

First selection of data was based on ICD-9-CM procedure and NIHDI procedure codes as proposed by Taylor et al., and by Jacques et al. This selection resulted in a total of 14 105 stays.[246, 309]

Table 7.19: Total knee arthroplasty: Primary data selection in MCD 2004

SELECTION 1 ICD-9-CM procedure code		SELECTION 2 NIHDI procedure code	
8154 8155	TOTAL KNEE REPLACEMENT REVISE KNEE REPLACEMENT	290264	Femoro-tibial arthroplasty with simple prosthesis
		290286	Femoro-tibial arthroplasty with articulated prosthesis
		293462	Removal of a total knee prosthesis and placement of a new total prosthesis
STAYS SELECTED =13 856		**STAYS SELECTED =13 762**	
TOTAL STAYS SELECTED= 14 105 (selection 1 OR selection 2)			

7.2.3 Definition of procedure

Table 7.20 shows that both ICD-9-CM and NIHDI procedures codes were recorded for almost 96% of the stays. There was thus a good correspondence between both classification systems.

Table 7.20: Total knee arthroplasty: Cross table ICD-9 and NIHDI procedures codes

	Number	Percent	Cumulative Freq	Cumulative Percent
ICD-9 only	343	2.43	343	2.43
NIHDI only	249	1.77	592	4.20
Both	13 513	95.80	14 105	100.00

Among the 13 513 stays with both an ICD-9 and a NIHDI code, there were 926 which related to a revision of the hip arthroplasty. These stays were not withheld.

Subsequently, we identified 12 752 stays with one total knee arthroplasty in which both the NIHDI procedure code 290286 as well as the ICD-9-CM procedure code 8151 were encoded. Stays with more than one TKR billed were then discarded (number=109) together with 190 stays with abnormalities in the data. After withdrawing the remaining 84 stays with a simple prosthesis (NIHDI code 290264), in order to obtain a homogeneous population, 12 369 stays could be further examined.

After this first selection step, we looked at the available clinical information. Major Diagnostic Category (MDC) 08 "Diseases & disorders of the musculoskeletal system" represents 99.5% of those 12 369 stays. Table 7.21 shows the APR-DRGs from these 12 312 stays. APR-DRG 302 comprises 99.6% of all stays in MDC 08.

Table 7.21 Total knee arthroplasty: All APR-DRGs from MDC 08-Diseases & disorders of the musculoskeletal system, with one TKR encoded with a NIHDI code AND an ICD-9 code

	COUNT	PERCENT
302-MAJOR JOINT & LIMB REATTACH PROC OF LOWER EXTREM EXC FOR TRAUMA / 8 P	12 260	99.58
300-BILATERAL & MULTIPLE MAJOR JOINT PROCS OF LOWER EXTREMITY / 8 P	11	0.09
301-MAJOR JOINT & LIMB REATTACH PROC OF LOWER EXTREMITY FOR TRAUMA / 8 P	41	0.33
TOTAL	12 312	100

Inside MDC 08, the knee replacements after trauma are excluded i.e. APR-DRG 301 representing 0.33% of the MDC 08. The last group to be excluded was APR-DRG 300 including the bilateral and multiple procedures and concerning 0.1% of MDC 08. After this step, 12 260 stays were withheld.

Table 7.22 shows that the majority (96.8%) of the stays in APR-DRG 302 had a principal diagnosis 715 "Osteoarthrosis & allied disorders". Twenty-one stays with principal diagnosis 996x "Complications peculiar to certain specified procedures" and 2 stays with V54x "Other orthopaedic aftercare" were rejected. This way, 12 237 stays remained.

Table 7.22 Total knee arthroplasty: List of all principal diagnostic ICD-9 codes in APR-DRG 302

	Number	Percent
715 Osteoarthrosis & allied disorders	11868	96.8
736 Other acquired deformities of limbs	134	1.09
733 Other disorders of bone & cartilage	104	0.85
717 Internal derangement of knee	32	0.26
716 Other & unspecified arthropathies	25	0.2
996 Complications peculiar to certain specified procedures	21	0.17
715 Rheumatoid arthritis & other inflammatory polyarthropathies	20	0.16
720 Other & unspecified disorders of joint	11	0.09
730 Osteomyelitis, periostitis, & other infections involving bone	11	0.09
170 Malignant neoplasm of bone & articular cartilage	8	0.07
718 Other derangement of joint	7	0.06
198 Secondary malignant neoplasm of other specified sites	2	0.02
712 Arthropathy associated with infections	2	0.02
722 Spondylosis & allied disorders	2	0.02
905 Late effects of musculoskeletal & connective tissue injuries	2	0.02
V54 Other orthopaedic aftercare	2	0.02
015 Tuberculosis of bones & joints	1	0.01
172 Malignant neoplasm of connective & other soft tissue	1	0.01
213 Benign neoplasm of bone & articular cartilage	1	0.01
696 Psoriasis & similar disorders	1	0.01
732 OSTEOCHONDROPATHIES	1	0.01
754 Certain congenital musculoskeletal deformities	1	0.01
805 Fracture of vertebral column without mention of spinal cord injury	1	0.01
812 Fracture of humerus	1	0.01
813 Fracture of radius & ulna	1	0.01
TOTAL	12260	100

The final step in the selection was to discard 962 patients who had undergone two TKR during two consecutive hospitalizations in the period 2004-2005. This was done to avoid confusion in identifying and allocating adverse events. The final population comprised only patients with one TKR in 2004-2005. The volume per hospital was thus derived from a strict selection of stays, assuming the consecutive steps in the selection were distributed evenly across all hospitals.

Results of definition: 11 017 stays with one elective total knee replacement were selected (for 11 017 patients).

7.2.4 Definition of volume

Table 7.23 shows that 114 hospitals performed 97 elective total knee arthroplasty on average. One hospital performed 557 procedures, the second in volume accounting for 424 procedures. Three quarters of the hospitals performed 120 elective TKR or less.

Table 7.23: Total knee arthroplasty: Summary measures of volume per hospital

Number	Minimum	Mean	25th Pctl	50th Pctl	75th Pctl	Maximum
114	4.0	96.6	46.0	78.0	120.0	557.0

Figure 7.8: Total knee arthroplasty: volume per hospital

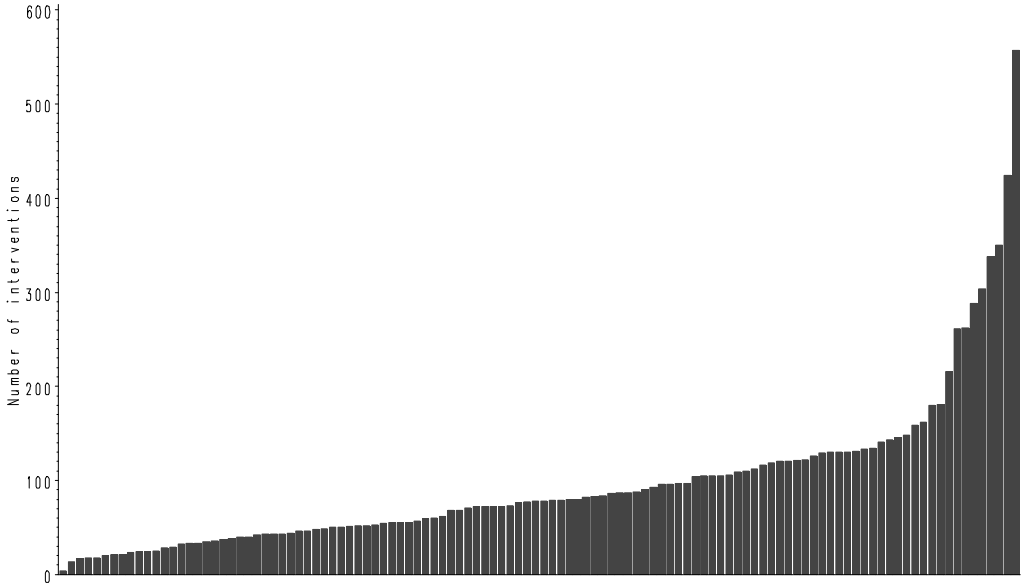


Table 7.24 gives information on the number of elective TKR per surgeon. For surgeons who work in several hospitals, a sum was made of all their TKR. In 2004, 488 orthopaedic surgeons performed at least one TKR. The mean number of TKR per surgeon is 23, while the median 12 is. The maximum was 290 elective TKR. Three quarters of surgeons performed 30 elective TKR or less.

Table 7.24: Total knee arthroplasty: Summary measures of volume per surgeon

Number	Minimum	Mean	std	25th Pctl	50th Pctl	75th Pctl	Maximum
488	1.0	22.6	28.6	5.0	12.0	30.0	290.0

Figure 7.9: Total knee arthroplasty: volume per surgeon

7.2.5 Definition of outcomes

Common outcome measures are mortality, loss of independence, loss of mobility and residual pain. Unfortunately, the latter three are not registered in the MCD. Complications associated with TKR can be divided in three kinds. Those associated with the risk of anaesthesia includes acute myocardial infarction. Those associated with immobilization include development of deep vein thrombosis, pulmonary embolism, pneumonia, and muscular atrophy; those associated with the surgical procedure include postoperative infection, mechanical malfunctions of the prosthesis and dislocation of the knee joint. These complications can give cause for a re-intervention, i.e. a revision knee replacement. Finally, in the long term, there is the problem of knee replacements progressively wearing out over the years. Patients who wear out their prosthesis will require a revision knee replacement surgery.

The literature review (see Table 2.1 on page 16) identified the following outcome measures for TKR: in-hospital mortality, in-hospital complication rate, 90-day complication rate and the 90-day revision rate. The revision rate was also studied at 36 months, but since MCD data were only available for 2004 and 2005, this long-term outcome measure was not retained in our analyses.[62]

7.2.5.1 In-hospital outcome

In-hospital outcome includes in-hospital mortality as well as complications after total knee arthroplasty. Information on in-hospital complications was retrieved from the MCD, on the basis of the following ICD-9-CM secondary diagnosis codes during index admission: [310]

- Deep wound infection (SSI): 996.66 "Infection and inflammatory reaction due to internal joint prosthesis"
- Pulmonary embolism: 415.1 "Pulmonary embolism and infarction"
- Deep venous thrombosis (DVT): 451.1x "Phlebitis and thrombophlebitis of deep vessels of lower extremities".

Table 7.25 shows that these four outcome measures are rare events. During the 11 017 elective TKR stays, 17 patients deceased during their hospitalization. The most frequent in-hospital complication was pulmonary embolism with 0.46% (0.34%, 0.61%). Similarly to what was done for total hip arthroplasty, no further analysis in terms of VOA was performed on in-hospital outcomes. Instead, all our attention was focused on events occurring within 90 days after the first index stay, and on the revision rate at 12 months.

Table 7.25: Total knee arthroplasty: In-hospital outcome (during index stay)

	Number total	number outcome	% outcome
Death	11017	17	0.15
Deep venous thrombosis	11017	45	0.41
Pulmonary embolism	11017	51	0.46
Deep wound infection	11017	7	0.06

7.2.5.2 Complications within 90 days

Among the 11 017 patients who underwent a total knee replacement in 2004, 6 633 patients (60.7%) were hospitalized at least once again in 2004 or 2005. For 1 875 patients (17%), the second admission took place within 90 days following the index admission.

The 90 days interval starts with the total knee replacement index admission. This period of time was chosen previously by Katz. et al.[325] The outcome measures are very similar to those for THR.

Information on the 90-day complication rate is retrieved from two sources:

1. the index admission: on the basis of the secondary diagnosis;
2. consecutive readmissions that occurred within 90 days from index admission: on the basis of the principal diagnosis or procedure.

The following ICD-9-CM diagnosis (as primary or secondary diagnosis) or procedure codes are used:

- Pulmonary embolism: 415.1 "Pulmonary embolism and infarction as principal diagnosis"
- Deep venous thrombosis (DVT): 451.1x "Phlebitis and thrombophlebitis of deep vessels of lower extremities"
- Deep wound infection: 996.66 "Infection and inflammatory reaction due to internal joint prosthesis"
- Mechanical complication: 996.4 "Mechanical complication of internal orthopaedic device, implant, & graft"
- Other complications: at least one ICD-9-CM diagnosis code indicating "Other complications of internal prosthetic device, implant, & graft":
  - 996.77 "Due to internal joint prosthesis"
  - or 996.78 "Due to other internal orthopaedic device, implant, & graft"
  - or 996.79 "Due to other internal prosthetic device, implant, & graft"
- Dislocation of knee within 90 days from initial admission: at least one of the following codes:
  - diagnosis 836.x "Dislocation of knee"
  - or procedure 79.76 "Closed reduction of dislocation of knee"
  - or procedure 79.86 "Open reduction of dislocation of knee"
- Acute myocardial infarction: 410 "Acute myocardial infarction"
- Pneumonia: at least one of the following ICD-9-CM diagnosis codes:
  - 480.x "Viral pneumonia"
  - 481.x "Pneumococcal pneumonia"
  - 482.x "Other bacterial pneumonia"
  - 483.x "Pneumonia due to other specified organism"
  - 485.x "Bronchopneumonia, organism unspecified"
  - 486.x "Pneumonia, organism unspecified"

Table 7.26 shows that the most frequent complication observed during hospitalization (initial or readmissions) within 90 days after admission was a complication due to internal joint prosthesis with 1.06% (0.88,% 1.27%), followed by 0.82% mechanical complication (0.66%, 1 %). As explained in the chapter on total hip arthroplasty (see page 174), there may be some overlap between categories, and some patients experienced different complications or are readmitted more than once. Therefore, a global outcome category was created to group all patients who encountered any of the above complications within 90 days after the initial TKR, with the exception of AMI and pneumonia. 3.20% (2.88%, 3.55%) of the patients (number = 353) encountered at least one of the above mentioned post-operative complications within 90 days after the initial total knee arthroplasty admission. AMI and pneumonia were left out of this global complication rate for two reasons: to be consistent with the definition for total hip arthroplasty and because AMI and pneumonia are more generic complications that are not particular for this orthopaedic intervention.

Table 7.26: Total knee arthroplasty: Outcomes measured during hospital (re)admission within 90 days of index admission

COMPLICATION WITHIN 90 DAYS	Number total	number outcome	% outcome
Deep venous thrombosis	11 017	46	0.42
Pulmonary embolism	11 017	69	0.63
Deep wound infection	11 017	34	0.31
Mechanical complication	11 017	90	0.82
Dislocation of knee	11 017	14	0.13
Other complication due to prosthesis	11 017	117	1.06
Any of the above complications *	**11 017**	**353**	**3.20**
Other complications:			
Acute myocardial infarction	11 017	22	0.20
Pneumonia	11 017	62	0.56

* a patient may have more than one complication, but these are counted only once

7.2.5.3 Revision rate at 12 months

Revision rates at 12 months were retrieved by identifying the readmissions with an ICD-9-CM procedure code 8155 or an NIHDI procedure code 293462 (see Table 7.19).

Table 7.27 shows that less than 1% of the patients needed a revision in the year after the operation (0.98%). Since we disposed of a 12-month follow-up for all patients, there was no need to estimate the rate of revision by the Kaplan-Meier method. Revision rate at 36 months, as is done in scientific literature, was not calculated because data were not available.

Table 7.27: Total knee arthroplasty: Revision rate at 12 months

	Number total	number outcome	% outcome	95% CI	
Readmission for revision at 12 months	11 017	108	0.98	0.80	1.18

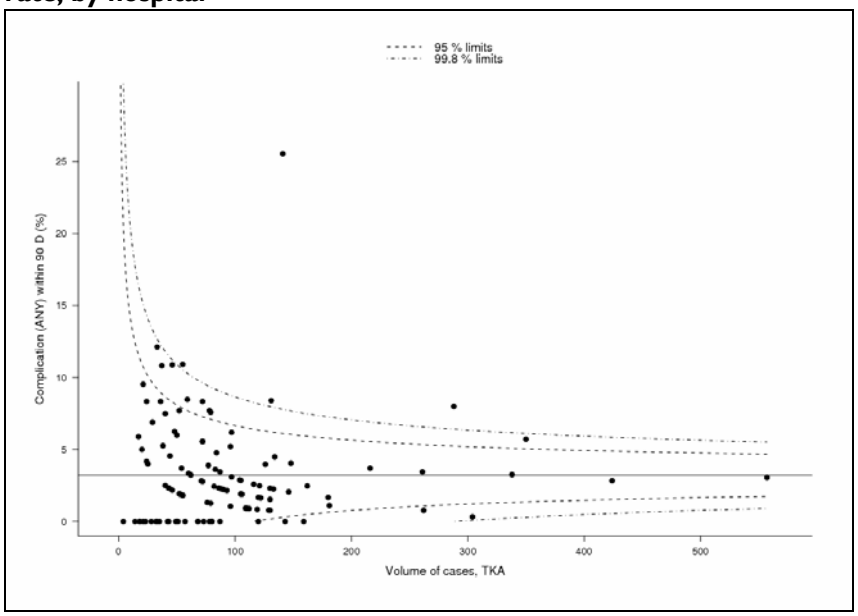
7.2.6 Volume outcome relationship

7.2.6.1 Analysis by hospital

90-DAY COMPLICATION (ANY) RATE, BY HOSPITAL

Figure 7.10 shows the association between the volume of interventions by hospital and the rate of complication within 90 days. The horizontal line represents the overall 90-day complication rate over all hospitals i.e. 3.2%. Three hospitals are located outside the variability limits and a fourth one is isolated far from the funnel plot. This latest hospital belonged to the third tertile with 141 TKR patients of whom 36 (25.5%) encountered a complication within 90 days (35 with a code 996.7x complication due to the internal device and one with a code 996.4 mechanical complication of the device). Such a manifest outlier would need to be further investigated by means of an audit.

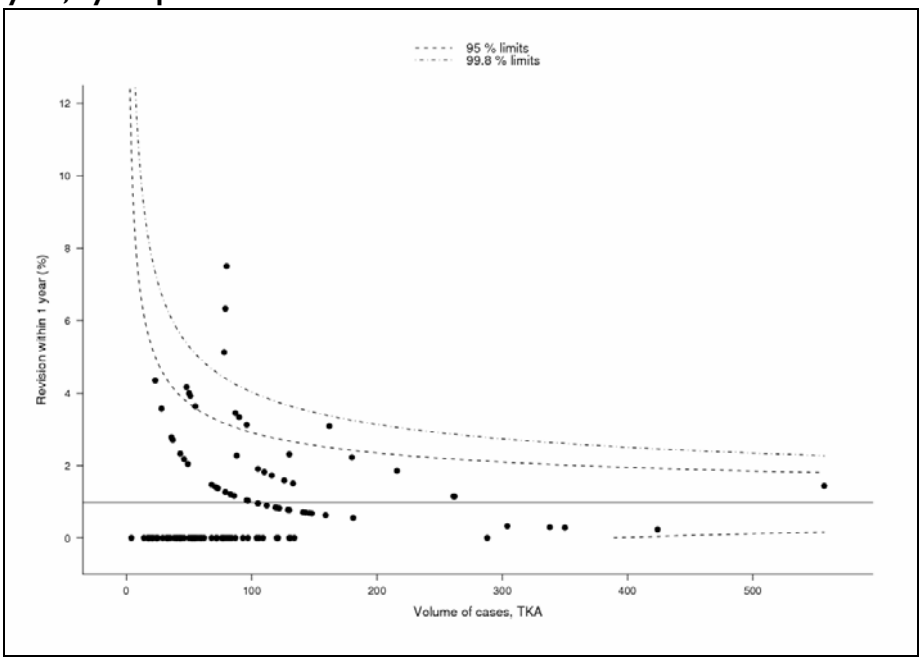
Figure 7.10: Total knee arthroplasty: Funnel plot of the 90-day complication rate, by hospital



REVISION RATE AT 1 YEAR, BY HOSPITAL

The funnel plot of the relationship between revision rate at 1 year and hospital volume is presented in Figure 7.11. Three medium-sized hospitals fall outside the limits of variability. They have a TKR volume of, respectively, 78, 79 or 80 interventions. The revision rate for their patients was 5.1%, 6.3%, 7.5%.

Figure 7.11: Total knee arthroplasty: Funnel plot of the revision rate at 1 year, by hospital



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

The differences in case-mix in function of the volume of total knee arthroplasty performed per hospital are presented in Table 7.28. More women were operated than men (72% versus 28%), especially in small hospitals (having performed 53 TKR or less in 2004). Mean age of patients was similar between the three groups. The percentage of high Charlson index score was the highest in medium-sized hospitals (53.6% versus 44.8% for small-sized hospitals and 47.8% for large sized ones). The revision rate within 1 year was highest when the index admission hospital was a medium-sized hospital. Remember that the highest rates were observed in 3 of those hospitals (see Figure 7.11). Complication rates within 90 days of small-sized hospitals were relatively higher than those of the two other categories of hospitals.

Table 7.28: Total knee arthroplasty: Case mix description by volume category of hospital

		Tertile based on volume			
		≤53 TKR	>53 and ≤104 TKR	>104 TKR	All
Number hospitals		38	38	38	114
Number stays		1322	2922	6773	11017
Male	%	29.8	26.4	28.2	27.9
Age	Mean	70.1	69.7	69.5	69.6
	Std	9.8	9.6	9.7	9.7
Charlson index score	Mean	0.5	0.5	0.5	0.5
	Std	1.0	1.0	1.0	1.0
Charlson ≥ 3	%	44.8	53.6	47.8	49.0
Principal diagnosis					
715 Osteoarthrosis	%	96.1	94.8	97.9	96.9
736 Other acquired deformities of limbs	%	0.4	3.1	0.4	1.1
Other principal diagnosis	%	3.6	2.1	1.7	2.0
Severity index (APR-DRG) for stays with principal diagnosis 715		1270	2771	6631	10672
1	%	45.3	54.8	48.0	49.4
2	%	43.9	38.2	44.9	43.1
3	%	10.0	6.2	6.6	6.9
4	%	0.8	0.8	0.5	0.6
Complication rate (any) within 90 days	%	3.7	3.1	3.1	3.2
Revision rate at 1 year	%	1.0	1.3	0.9	1.0

RELATION BETWEEN HOSPITAL VOLUME AND 90-DAY COMPLICATION (ANY)

Table 7.29 shows the results in terms of effects of the logistic regression aiming at explaining the 90-day complication rate (any). Without adjusting for case-mix, a 10% volume increase is associated with a neglible decrease in the odds of complication of -0.63% (-3.37%, 2.11%). Males present better odds than female patients and the Charlson score does influence the complication rate.

Table 7.29: Total knee arthroplasty: Correlation-corrected logistic regression (GEE) estimates of determinants of 90-day complication

Model without adjustment for case mix	Effect[1]	95% CI	
Hospital volume (increase of 10%)	-0.63	-3.37	2.11
Model with adjustment for case mix	Effect1	95% CI	
Hospital volume (increase of 10%)	-0.36	-2.99	2.28
	Odds Ratio	95% CI	
Sex (male vs. female)	0.73	0.56	0.94
Age (increase of 1 year)	0.99	0.97	1.00
Charlson score (increase of 1 category)	1.39	1.20	1.60
715 vs. other PDX	0.60	0.36	1.00

[1] Effect of 10% increase in volume of surgeon on the odds of complication within 90 days

RELATION BETWEEN HOSPITAL VOLUME AND REVISION RATE AT 1 YEAR

Results of logistic regression presented in Table 7.30 show that the effect of a volume increase of the procedures of a hospital on the revision rate within 1 year is very small and non statistically significant neither before (-0.81%; -3.52%, 1.9%), nor after adjustment for the case-mix differences (-1%; -3.74, 1.75). There is thus no indication of a relationship between hospital volume and risk for revision after TKR.

Table 7.30: Total knee arthroplasty: Results of regression: relative effects of 10% increase volume of hospital on revision rate within 1 year

Model without adjustment for case mix	Effect[1]	95% CI	
Hospital volume (increase of 10%)	-0.81	-3.52	1.90
Model with adjustment for case mix	Effect1	95% CI	
Hospital volume (increase of 10%)	-1.00	-3.74	1.75
	Odds Ratio	95% CI	
Sex (male vs. female)	0.55	0.34	0.9
Age (increase of 1 year)	0.96	0.94	0.98
Charlson score (increase of 1 category)	1.21	0.96	1.51
715 vs. other PDX	2.12	0.53	8.39

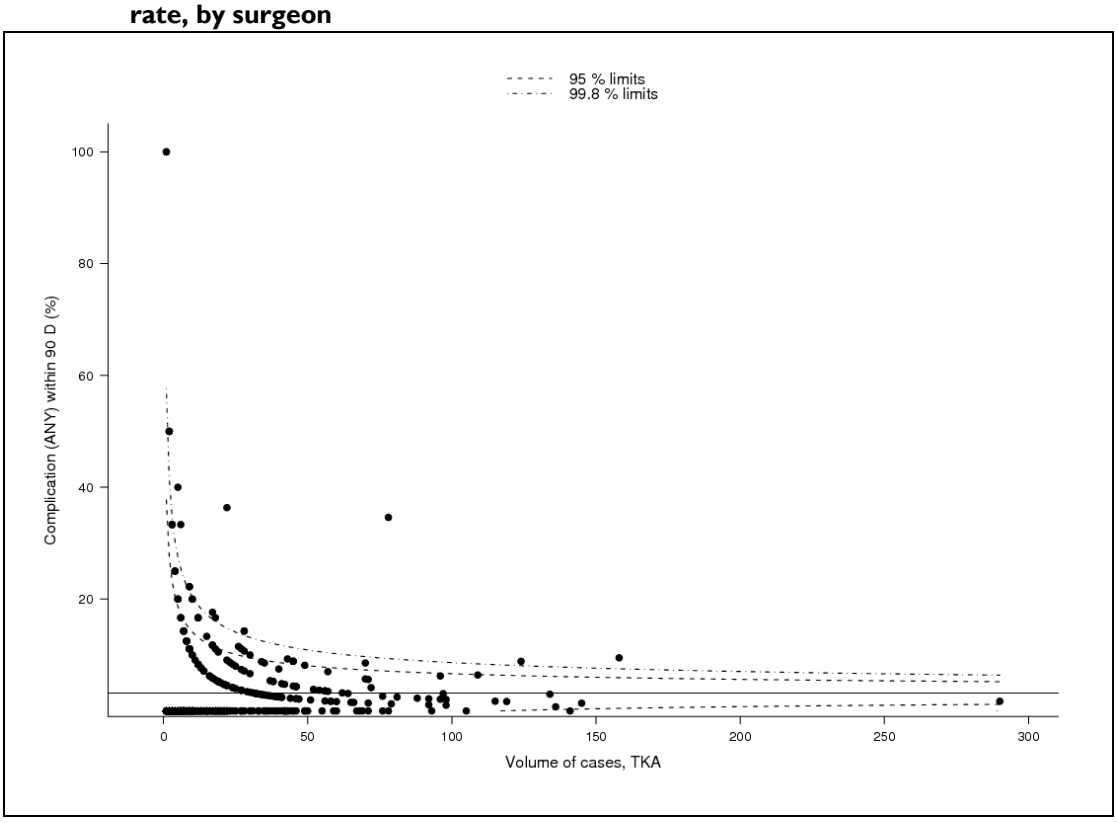
[1] Effect of 10% increase in volume of surgeon on the odds of revision within 1 year

7.2.6.2 Analysis by surgeon

90-DAY COMPLICATION (ANY) RATE, BY SURGEON

The funnel plot distribution of the relationship between the complication rate within 90 days and the volume of interventions by surgeon is presented in Figure 7.12. About 9 surgeons were located outside the 95% limits and may therefore be considered as outliers that need further investigation. The most preoccupying points were two surgeons with a 34.6% (27/78) and 36% (8/22) 90-day complication rate. The surgeon with the highest experience was located inside the limits, below the national average.

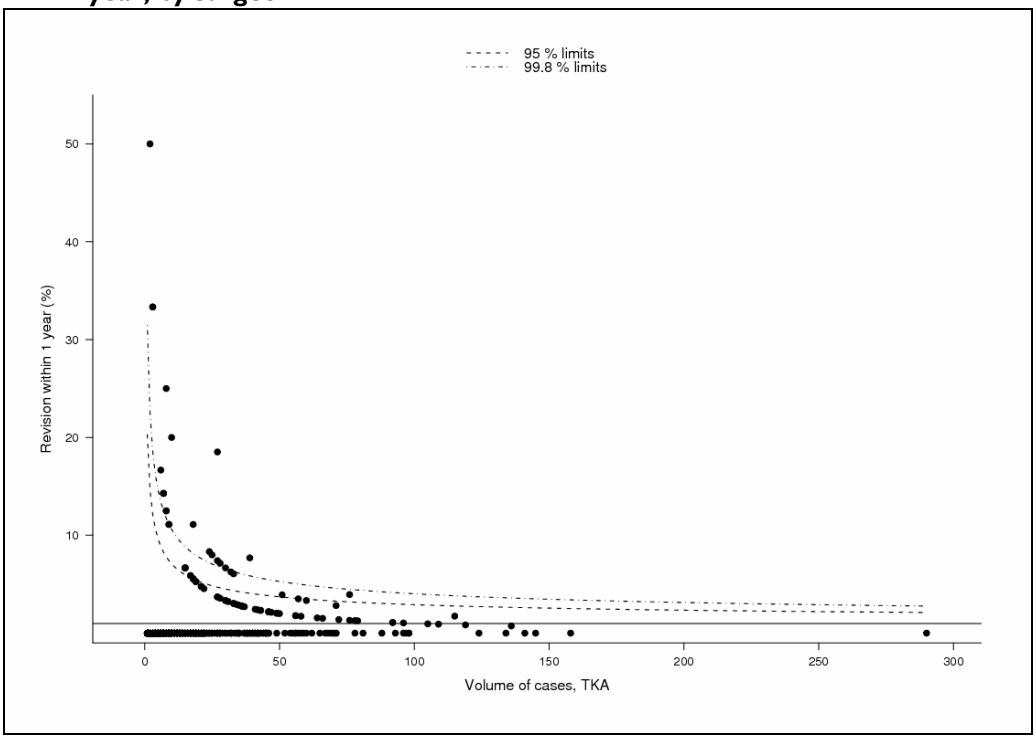
Figure 7.12: Total knee arthroplasty: Funnel plot of the 90-day complication rate, by surgeon



REVISION RATE AT 1 YEAR, BY SURGEON

Figure 7.13 presents the association between volume of procedures and revision rate within 1 year and shows that a few surgeons with less than 50 TKR in 2004 have higher rates than what could be expected from random variation; the furthest point being located at 27 interventions and 18.5% revision rate.

Figure 7.13: Total knee arthroplasty: Funnel plot of the revision rate at 1 year, by surgeon



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE SURGEON

The distributions of patient characteristics according to the volume of the operating surgeon are presented in Table 7.31. Complication rates and revision rates are 3.9% and 1.1% for the smallest volume surgeons and 3.1% and 0.9% for the highest volume surgeons. The percentage of male patients is lower for the small volume surgeons then for the two other categories. Low-volume surgeons also have patients with a higher Charlson score, but the severity of illness did not seem higher than for the two other categories of surgeons.

Table 7.31: Total knee arthroplasty: Case mix description by volume category of surgeon

		Tertile based on volume			
		≤7 TKR	7 TKR > and ≤21 TKR	>21 TKR	All
Number surgeons		177	149	162	488
Number stays		611	2062	8344	11017
Male	%	25.5	28.1	28.1	27.9
Age	Mean	68.3	69.5	69.7	69.6
	Std	11.6	9.8	9.4	9.7
Charlson index score	Mean	0.5	0.5	0.5	0.5
	Std	1.1	1.0	0.9	1.0
Charlson ≥ 3	%	5.2	4.8	4.4	4.5
Principal diagnosis					
715 Osteoarthrosis	%	96.7	95.9	97.1	96.9
736 Other acquired deformities of limbs	%	0.2	0.7	1.3	1.1
Other principal diagnosis	%	3.1	3.4	1.6	2.0
Severity index (APR-DRG) for stays with principal diagnosis 715		591	1977	8104	10672
1	%	48.6	47.8	49.9	49.4
2	%	43.3	43.6	42.9	43.1
3	%	7.8	7.8	6.7	6.9
4	%	0.3	0.9	0.6	0.6
Complication rate (any) within 90 days	%	3.9	3.2	3.1	3.2
Revision within 1 year	%	1.1	1.1	0.9	1.0

RELATION BETWEEN SURGEON VOLUME AND 90-DAY COMPLICATION (ANY) RATE

As presented in Table 7.32, the association between a 10% increase in surgeon volume and complication rate was an effect of -2.29% (-3.96%, -0.62%), without adjusting for case-mix differences. Case-mix has no impact of since the effect sustains after such case-mix adjustment although the decrease is slightly less important (-2.16; -3.82%, -0.5%). Both effects were statistically significant. Similar to the results by hospital, odds for complication are worse for women.

Table 7.32: Total knee arthroplasty: Results of regression: relative effects of 10% increase volume of surgeon on 90-day complication (any)

Model without adjustment for case mix	Effect[1]	95% CI	
Surgeon volume (increase of 10%)	-2.29	-3.96	-0.62
Model with adjustment for case mix	Effect[1]	95% CI	
Surgeon volume (increase of 10%)	-2.16	-3.82	-0.50
	Odds Ratio	95% CI	
Sex (male vs. female)	0.55	0.34	0.88
Age (increase of 1 year)	0.96	0.94	0.98
Charlson score (increase of 1 category)	1.19	0.89	1.60
715 vs. other PDX	2.34	0.56	9.72

[1] Effect of 10% increase in volume of surgeon on the odds of complication within 90 days

RELATION BETWEEN SURGEON VOLUME AND REVISION RATE AT 1 YEAR

Logistic regression results for the revision rate are presented in Table 7.33. There was an effect of -3.72% (-5.94%, -1.5%) on the revision rate when the volume of surgeon increased by 10%. When we adjusted for case-mix, this effect was slightly reduced to -3.65%, but it remained statistically significant.

Table 7.33: Total knee arthroplasty: Results of regression: relative effects of 10% increase volume of surgeon on revision rate within 1 year

Model without adjustment for case mix	Effect[1]	95% CI	
Surgeon volume (increase of 10%)	-3.72	-5.94	-1.50
Model with adjustment for case mix	**Effect1**	**95% CI**	
Surgeon volume (increase of 10%)	-3.65	-5.86	-1.44
	Odds Ratio	95% CI	
Sex (male vs. female)	0.56	0.35	0.91
Age (increase of 1 year)	0.96	0.94	0.98
Charlson score (increase of 1 category)	1.14	0.83	1.55
715 vs. other PDX	2.31	0.55	9.75

[1] Effect of 10% increase in volume of surgeon on the odds of revision within 1 year

7.2.7 Discussion

7.2.7.1 External validation of the definition of the procedure

We analyzed about 11 000 stays out of 14 000 stays (revisions included) identified in the MFD or MCD with a code of knee arthroplasty (see Table 7.19). When we take into account a yearly increase in numbers, our number corresponds well with the annual number of stays observed in the MCD between 1997 and 2002, which was around 10 000 (revisions included) in the KCE report by Jacques et al..[246]

7.2.7.2 Summarized results of literature review

The systematic literature search identified seven systematic reviews in which the VOA for TKR was studied.[1, 5, 59, 60, 62, 64, 68] In all, these systematic reviews identified 16 primary studies of which 5 were published between 2002 and 2004,[325-329] and 10 between 1995 and 1999.[309, 317, 321, 330-336] This number was considered sufficient for the discussion.

The best systematic review is the meta-analysis by Stengel et al. because it is quite recent (2004) and of good quality (see Appendix 7).[68] Stengel et al. meta-analysed data from 5 articles of which 4 included risk-adjustment.[309, 317, 326, 328, 329] They analyzed the effect of hospital volume on in-hospital mortality, 90-day mortality, overall complication rate, infection rate and DVT rate. The results of these analyses are described in Appendix 10, and are mentioned in the discussion below.

On the basis of all seven systematic reviews, it was concluded in Chapter 2 (see Table 2.2 on page 19) that there is an inverse relation between hospital volume and in-hospital mortality. For post-operative complications, on the other hand, there were conflicting results in relation to the association with hospital volume. The term "conflicting results" is used because there is a mix of primary studies that indicate a positive relation with volume and others that indicate a negative relation with volume. Since there was only one low-quality systematic review which studied the revision rate, these results could not be summarized in the evidence table in Appendix 10.[62]

7.2.7.3 *Comparative analysis of literature and Belgian data*

DEFINITION OF VOLUME

As shown in Table 7.23, Belgian hospitals had a mean annual elective TKR volume of 97 in 2004; median is 78; 75th Percentile is 120. Table 7.24 shows that Belgian orthopaedic surgeons performed on average 23 elective TKR in 2004; median is 12; 75th Percentile is 30 TKR per year. These summary measures are very similar to those for elective total hip replacement (see page 49). When we compare the highest volume hospitals for elective THR (Figure 5.1) with those for elective TKR (Figure 6.2), the same hospital IDs turn up on the horizontal axis. This means that Belgian hospitals with a high annual volume of THR, often also have a high volume of elective TKR.

In the selected primary studies and the systematic review, we did not find any information on the annual TKR volume of hospitals or surgeons in other countries.

OUTCOME

Again we have to acknowledge the fact that the Belgian MCD do not capture clinically relevant outcome measures such as loss of dependence, loss of mobility or residual pain.

Table 7.34: Total knee arthroplasty: complication rate: comparison of Volume-Outcome KCE study with scientific literature

IN-HOSPITAL COMPLICATIONS	KCE study Volume outcome	Hervey 2003	Coyte 1999	Kreder 2003	Katz 2004
Complications during index admission					
Death	0.15%	0.21%		0.4%	
Deep venous thrombosis	0.41%	0.48%			
Pulmonary embolism	0.46%	0.37%			
Deep wound infection	0.06%	0.25%			
Complications within 90 days from index admission					
Death				0.5%	0.6%
Deep venous thrombosis	0.42%				
Pulmonary embolism	0.63%				0.8%
Deep wound infection	0.31%				0.4%
AMI	0.20%				0.8%
Pneumonia	0.56%				1.4%
Complications within 1 year (or more) after index admission					
Revision within 1 year	0.98%			0.8%	
Revisions within 500 days (1.36 year)			0.2% (*) 1.6% (*)		

(*) Coyte et al used two methods to estimate the revision rate: based on a group of patients with the longest time to revision and based on a group with the shortest time.

In-hospital outcome after total knee replacement

In-hospital outcome in the KCE study was first compared with the study by Hervey on 1997 data.[328] See Table 7.34. Hervey's complication rates were very similar to the ones that were found in the MCD data except for deep wound infection. This difference can probably be explained by the use of different codes. Hervey et al. identified the infections by means of ICD-9-CM codes 998.59 "Postoperative wound infection" and 686.9 "Postoperative wound infection knee (skin)" while we used 996.66 "Infection and inflammatory reaction due to internal joint prosthesis". The latter code was an exclusion criterion in the Hervey study. Kreder et al. reported an in-hospital mortality of 0.4%, but this is perhaps related to the fact that it concerns older data i.e. TKR performed between 1993 and 1996.[329]

90-day complication rate after total knee replacement

In Table 7.34 we compare the KCE study's 90-day complication rates with those published by Katz et al. on elective TKR performed in 2000.[325] The rates for pulmonary embolism and deep wound infection are very similar. For AMI and pneumonia, on the other hand, Katz obtained substantially higher rates. The explanation for these differences is again related to the use of different codes. For pneumonia, the KCE study used the ICD-9-CM codes 480 to 486 while Katz selected a larger range of codes (according to the AHRQ algorithms). For AMI, the KCE used code 410.xx while Katz's AHRQ algorithm identified only the stays with 410.x1. On the other hand, the KCE only considered the principal diagnosis of re-admissions while Katz also considered the secondary diagnoses of AMI.

Revision rate at 12 months after total knee replacement

Table 7.34 shows that the 1-year revision rate observed in our study (0.98%) is very similar to the one found by Kreder (0.80%).[329] Coyte and colleagues found a 500-days revision rate between 1.6% and 2.0% (according to the method used).[331] Nevertheless, the reader has to take into account the fact that Coyte used data on TKR performed between 1984 and 1991.

Patient Case Mix

Table 7.28 and Table 7.31 showed that low-volume (≤53 TKR/year) and medium-volume hospitals (between 53 and 104 TKR/year) and low-volume surgeons (≤7 TKR/year) have higher risk patients based on a higher Charlson score and/or a higher severity of illness. Kreder and colleagues found that medium-volume surgeons and low-volume hospitals had a slightly higher percentage of patients with more than one co-morbidity.[329]

Volume Outcome Association

Relation between hospital or surgeon volume and in-hospital mortality

Because in-hospital death was such a rare event in the MCD (0.15%, see

Table 7.25), it was decided not to analyze the VOA for this outcome measure. Nevertheless, it is interesting to know that other studies managed to prove an inverse relation. Stengel et al. meta-analysed data from studies by Culler, Taylor, Kreder, Feinglass and Hervey.[68, 309, 317, 326, 328, 329] They found a small, but statistically significant association between hospital TKR volume and in-hospital mortality; between 2 551 and 821 TKR must be performed by high-volume rather than by low-volume hospitals to prevent one extra death.[68]

Relation between hospital or surgeon volume and 90-day complication (any) rate

After adjustment for age, gender, Charlson score and principal diagnosis, higher hospital volume was associated with a lower rate of complications within 90 days after the index TKR, but the trend did not reach significance (see Table 7.29). Higher surgeon volume, on the other hand, was statistically significantly associated with a lower 90-day complication rate after elective TKR (see Table 7.32).

Contrary to what we concluded on the basis of the systematic reviews - that results for a VOA between post-operative complications and hospital volume were conflicting – the primary study by Katz et al. seems quite confident about the existence of such an association. Katz and colleagues assessed the relationship between surgeon and hospital procedure volume and mortality, deep wound infection (requiring surgical debridement or removal of the prosthesis), pulmonary embolism, myocardial infarction or pneumonia in the first ninety days postoperatively.[325] This assessment was done with claims data of Medicare patients who underwent an elective TKR in 2000. Analyses were adjusted for age, gender, arthritis diagnosis, a proxy for low income and the Charlson co-morbidity index. In addition, analyses of hospital volume were adjusted for surgeon volume and vice versa.

The analyses showed that patients who had a TKR in hospitals with more than 200 TKR per year had a lower risk of pneumonia (odds ratio 0.65; 99% CI 0.47 to 0.90) than those managed in hospitals with an annual volume of 25 TKR or fewer. These patients also had a lower risk of any of the adverse outcomes examined, i.e. mortality, deep wound infection, pulmonary embolism, myocardial infarction or pneumonia (odds ratio 0.74; 99% CI 0.60 to 0.90). Similarly, patients who were operated on by surgeons who performed more than 50 TKR per year, had a lower risk of pneumonia (odds ratio 0.72; 99% CI 0.54 to 0.95) and any of the adverse events (odds ratio 0.81; 99% CI 0.68 to 0.99) in comparison with patients of surgeons with fewer than 13 TKR per year.

Stengel et al. were less successful at proving a statistically significant VOA between TKR and the complication rate.[68] The risk of dying within 90 days postoperatively is less in higher-volume hospitals, as is the risk on a postoperative infection, but none of these trends was statistically significant. On the other hand, Stengel found that patients operated in higher volume hospitals were more at risk of a surgery-related complication, although this effect was also not statistically significant. Only for deep venous thrombosis, the meta-analysis showed a statistically significant association with the hospital volume.

RELATION BETWEEN HOSPITAL OR SURGEON VOLUME AND REVISION RATE AT 12 MONTHS

After adjustment for age, gender, Charlson score and principal diagnosis, higher hospital volume was associated with a lower 1-year revision rate after an elective TKR, but the trend did not reach significance (see Table 7.30). Higher surgeon volume, on the other hand, was statistically significantly associated with a lower 1-year revision rate (see Table 7.33).

Coyte and associates found a lower revision rate in community hospitals compared with teaching hospitals in Ontario.[331] The hospital volume as such was not taken into account.

Kreder et al. analyzed a cohort of Canadian patients that had undergone an elective TKR between 1993 and 1996.[329] Higher revision rates at 1 and 3 years were found to be significantly associated with lower patient age and low hospital volume. Hospitals with fewer than 48 TKR per year were 2.2 times (1.1, 4.5) more likely to require revision within 1 year of their index admission than hospitals performing more than 113 TKR annually. Contrary to what we found in the Belgian MCD, surgeon volume was not significantly associated with the incidence of revision at one year.

RELATION BETWEEN HOSPITAL OR SURGEON VOLUME AND POSTOPERATIVE FUNCTIONAL STATUS AND PAIN RELIEF

Unfortunately, Belgian MCD do not provide the outcomes of greatest interest to patients i.e. pain relief and improvement in functional status.

Heck and associates performed a prospective, observational cohort investigation among 291 patients who had a TKR in the state of Indiana (USA), in 1992.[335] They showed that patients in higher-volume hospitals (>50 TKR/year) were more likely to have an improvement in functional status (measured by the Short Form 36 Physical Composite Score) following TKR.

Key points on volume outcome association for elective total knee replacement/arthroplasty (TKR)

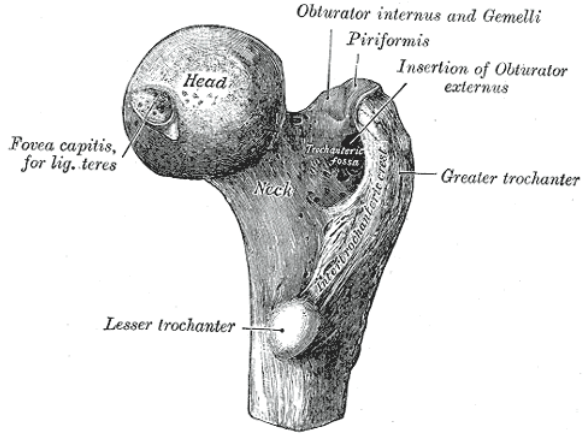
- **A total of 11 017 patients were hospitalized in 2004 for a total knee replacement. Revisions, bilateral procedures and procedures for trauma were not included in this selection.**
- **These interventions were performed in 114 centres by 488 surgeons. Annual mean number of TKR was approximately 100 per centre, and 23 per surgeon. These volumes are very similar to those of elective TKR.**
- **In-hospital mortality is virtually inexistent after elective TKR, and in-hospital complications are also very rare (or under reported in the MCD). Following outcomes were analysed: complications requiring a hospitalisation within 90 days of intervention (3.2% of all patients), and revision of the prosthesis at 12 months (1.0% of all patients). Logistic regression models were fitted to assess association between these outcomes and centre or surgeon volume. Factors included in the models were: age, sex, Charlson score and principal diagnosis (ostheoarthrosis versus others diagnoses).**
- **The literature review found conflicting results (i.e. mix of negative and positive VOA) in relation to the association between hospital volume and complication rate.**
- **Complication rate at 90 days was 3.7% for centres with less than 53 TKR/year, and 3.1 % for higher volume hospitals. These effects were too small to draw any conclusions on the effect of hospital volume.**
- **The association between surgeon volume and likelihood of complications at 90 days was small (and statistically significant with or without adjustment for case-mix): 3.9% for surgeons with less or 7 TKR/year, 3.2% for surgeons between 7 and 21, and 3.1% for surgeons with more than 21 TKR/year.**
- **The literature review did not succeed in identifying good quality systematic reviews that studied revision rate after TKR.**
- **There was no association between centre volume and 1-year revision rate.**
- **A small (and statistically significant with or without adjustment for case mix) decrease was observed in revision rates at 12 months from small or medium volume surgeons (less or 21 TKR/year, 1.1%) in comparison with high volume surgeons (> 21 TKR/year, 0.9%).**
- **Unfortunately, the MCD do not provide information on the outcomes of greatest interest to patients such as loss of independence, loss of mobility or residual pain.**
- **POSSUM (Physiological and Operative Severity Score for enUmeration of Mortality and morbidity) would be better for risk adjustment in orthopaedic surgery than the applied Charlson score. However, these clinical parameters are not encoded in the MCD.**

7.3 HIP FRACTURE SURGERY (HFS)

7.3.1 General description of procedure

Proximal femur fractures are located at the upper (i.e. proximal) extremity of the femur (i.e. thigh bone). This upper extremity presents a head, a neck and a greater and a lesser trochanter as is shown in Figure 6.1.

Figure 7.14: The proximal extremity of the femur



Source: Gray's Anatomy – Public domain

The majority of proximal femur fractures or hip fractures occur in elderly populations with an average age of 80 years and are a result of falling from standing height. Proximal femur fractures can be divided into intracapsular and extracapsular fractures. Intracapsular fractures are those occurring proximal to the femoral attachment of the hip joint capsule; these include femoral head and neck fractures. Extracapsular fractures are those occurring distal to the hip joint capsule; these include trochanteric, intertrochanteric, and subtrochanteric fractures. The most common fractures are the femoral neck fractures and the intertrochanteric fractures, which, respectively, account for approximately 40 and 50% of proximal femur fractures.[337]

Proximal femur fractures are almost always treated with surgery although the surgical method often differs by the type of fracture. Patients 60 years or younger with undisplaced or minimally displaced intracapsular fractures are most often treated using internal fixation (i.e. hip pinning). For older patients, there is still debate between fixation and arthroplasty. Extracapsular fractures, on the other hand, are usually treated using a sliding nail fixation.[337]

7.3.2 Primary data selection in MCD 2004

First selection of data was based on ICD-9-CM diagnosis codes as proposed by Hamilton et al., and, on the other hand, on ICD-9 procedure codes and NIHDI procedure codes as proposed by the external experts.[338] This selection resulted in a total of 31 339 stays.

Table 7.35: Hip fracture surgery: Primary data selection in MCD 2004

SELECTION 1		SELECTION 2		SELECTION 3	
ICD-9-CM diagnosis code		ICD-9-CM procedure code		NIHDI procedure code	
820xx	Fracture neck of femur*	7905	Closed reduction of fracture without internal fixation (hip)	289343	Surgical treatment of a subtrochanteric femoral fracture
		7915	Closed reduction of fracture with internal fixation (hip)		
		7925	Open reduction of fracture without internal fixation (hip)	289365	Surgical treatment of a per- or intertrochanteric femoral fracture
		7935	Open reduction of fracture with internal fixation (hip)		
				289380	Surgical treatment of a femoral neck fracture with hip pinning
				289402	Surgical treatment of a femoral neck fracture with a prosthesis
				289085	Hip arthroplasty with a total prosthesis (acetabulum and femoral head)
STAYS SELECTED =15 236		**STAYS SELECTED = 10 063**		**STAYS SELECTED =25 679**	
TOTAL STAYS SELECTED= 31 339 (selection 1 OR selection 2 OR selection 3)					

7.3.3 Definition of procedure

Since the NIHDI procedure code 289085 is not exclusively used for hip fracture surgery, it is really necessary to select only those hospital stays with a combination of the ICD-9-CM diagnosis code for femoral neck fracture (820.xx) AND one of the NIHDI procedure codes.

The ICD-9-CM procedure codes are of lesser importance for the selection of the procedures. Anyhow, for reasons of homogeneity, the closed reductions of hip fractures will not be considered in the analysis.

Table 7.36 confirms our suspicion that the primary data selection would also contain hospital stays for other reasons than a hip fracture: 45% of stays have a NIHDI procedure but no diagnosis of fracture of neck of femur; for 0.6% of stays femoral neck fracture was encoded only as secondary diagnosis.

It is therefore necessary to combine both criteria. This results in 10 690 stays (3 524 + 7 166 stays) with a combination of the ICD-9-CM diagnosis code for femoral neck fracture (820.xx) as principal diagnosis AND one of the NIHDI procedure codes.

Table 7.36: Hip fracture surgery: Presence of ICD-9 diagnosis code 820 in primary or secondary diagnosis AND ICD-9 and NIHDI procedures codes

Number Percent	ICD-9 proc only	NIHDI only	both	none	
Not in diagnoses	1654 5.28	13965 44.56	484 1.54	0 0.00	16103 51.38
Secondary diagnoses	65 0.21	182 0.58	358 1.14	662 2.11	1267 4.04
Principal diagnosis	336 1.07	**3524** **11.24**	**7166** **22.87**	2943 9.39	13969 44.57
Total	2055 6.56	17671 56.39	8008 25.55	3605 11.50	31339 100.00

Table 7.37 shows that 99% of these 10 690 stays are situated in MDC 08 (musculoskeletal system).

Table 7.37: Hip fracture surgery: List of the major diagnostic categories (MDC)

	COUNT	PERCENT
08 Diseases and Disorders of the Musculoskeletal System & Connective Tissue	10545	98.64
25 Multiple Significant Trauma	130	1.22
00 Residual Group	15	0.14
Total	10690	100.0

All APR-DRGs of MDC 08, as presented in Table 7.38, are surgical of nature. A surgical APR-DRG for trauma was assigned in most of the cases: APR-DRG 308 in 68% of the stays and APR-DRG 301 in another 31%. Less than 1% belongs to another APR-DRG.

Table 7.38: Hip fracture surgery: All APR-DRGs from MDC 08-Diseases & disorders of the musculoskeletal system

	COUNT	PERCENT
308-HIP & FEMUR PROCEDURES EXCEPT MAJOR JOINT FOR TRAUMA / 8 - P	7194	68.22
301-MAJOR JOINT & LIMB REATTACH PROC OF LOWER EXTREMITY FOR TRAUMA / 8 - P	3264	30.95
300-BILATERAL & MULTIPLE MAJOR JOINT PROCS OF LOWER EXTREMITY / 8 - P	51	0.48
305-AMPUTATION FOR MUSCULOSKELET SYSTEM & CONN TISSUE DISORDERS / 8 - P	7	0.07
320-OTHER MUSCULOSKELETAL SYSTEM & CONNECTIVE TISSUE PROCEDURES / 8 - P	7	0.07
313-KNEE & LOWER LEG PROCEDURES EXCEPT FOOT / 8 - P	6	0.06
315-SHOULDER, ELBOW & FOREARM PROCEDURES / 8 - P	5	0.05
316-HAND & WRIST PROCEDURES / 8 - P	3	0.03
306-MAJOR JOINT & LIMB REATTACHMENT PROCEDURES OF UPPER EXTREMITY / 8 - P	2	0.02
304-DORSAL & LUMBAR FUSION PROC EXCEPT FOR CURVATURE OF BACK / 8 - P	1	0.01
310-BACK & NECK PROCEDURES EXCEPT DORSAL & LUMBAR FUSION / 8 - P	1	0.01
312-SKIN GRFT & WND DEBRID EXC OPN WND, FOR MS & CONN TIS DIS, EXC HAND / 8 - P	1	0.01
314-FOOT PROCEDURES / 8 - P	1	0.01
318-REMOVAL OF INTERNAL FIXATION DEVICE / 8 - P	1	0.01
950-EXTENSIVE PROCEDURE UNRELATED TO PRINCIPAL DIAGNOSIS / 0 - P	1	0.01
	10545	100.0

For reasons of homogeneity, it was decided to discard bilateral procedures (APR-DRG 300 = 51 stays) and amputations (APR-DRG 305 = 7 stays). Subsequently, another 33 stays were discarded due to data irregularities.

A last filter was applied on the 10 454 remaining stays in order to exclude patients who were selected more than once (number = 215) or who were readmitted for hip fracture surgery within 90 days in the period 2004-2005 (number = 305). This was done to avoid confusion in identifying and allocating adverse events. The final population included 9 934 stays.

Results of definition: 9 934 stays with a proximal femur fracture were selected.

7.3.4 Definition of volume

Table 7.39 shows that 113 hospitals surgically treated 88 hip fractures on average, in 2004. One hospital performed 225 procedures, the second in volume accounting for 222 procedures. Three quarters of the hospitals treated 116 hip fractures or less.

Table 7.39: Hip fracture surgery: Summary measures of volume per hospital

Number	Minimum	Mean	25th Pctl	50th Pctl	75th Pctl	Maximum
113	7.0	87.9	52.0	78.0	116.0	225.0

Figure 7.15: Hip fracture surgery: volume per hospital

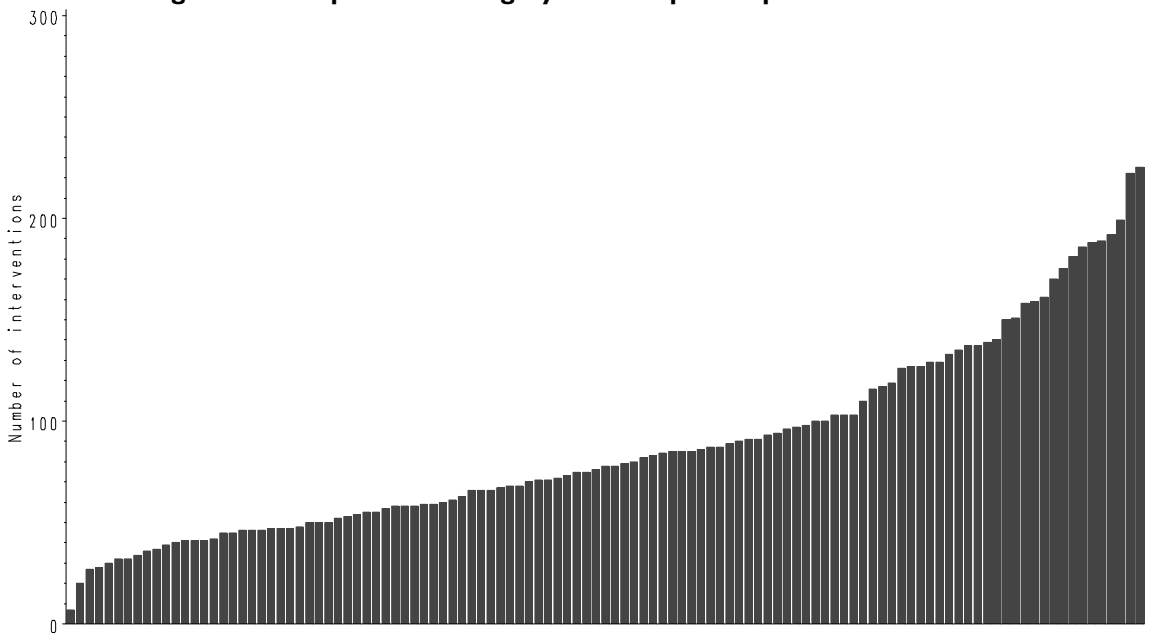


Table 7.40 gives information on the number of surgically treated hip fractures per surgeon. For surgeons who work in several hospitals, a sum was made of all their hip fracture surgeries. On average, a surgeon surgically treated 15 hip fractures in 2004. The maximum was 91 surgically treated hip fractures. Three quarters of surgeons treated 21 hip fractures or less.

Table 7.40: Hip fracture surgery: Summary measures of volume per surgeon

Number	Mean	Std Dev	Minimum	25th Pctl	Median	75th Pctl	Maximum
675	14.72	11.98	1	6	12	21	91

Figure 7.16: Hip fracture surgery: Volume per surgeon

7.3.5 Definition of outcomes

Common outcome measures are mortality, loss of independence, loss of mobility and residual pain. Complications associated with hip fracture surgery can be divided in two kinds. Those associated with immobilization include development of deep vein thrombosis, pulmonary embolism, pneumonia, and muscular atrophy; those associated with the surgical procedure include postoperative infection, loss of fixation and non-union. The latter complications can give cause for a re-intervention.[337]

The literature review (see Table 2.1 on page 16) only identified in-hospital mortality as outcome measure for proximal femur fracture.

With 645 deaths on a total of 9 934 cases (Table 7.41), in-hospital mortality after hip fracture surgery was 6.49% (6.02%, 7.00%).

Table 7.41: Hip fracture surgery: In-hospital mortality (index stay)

	Number total	number outcome	% outcome	95% CI	
Death	9 934	645	6.49	6.02	7.00

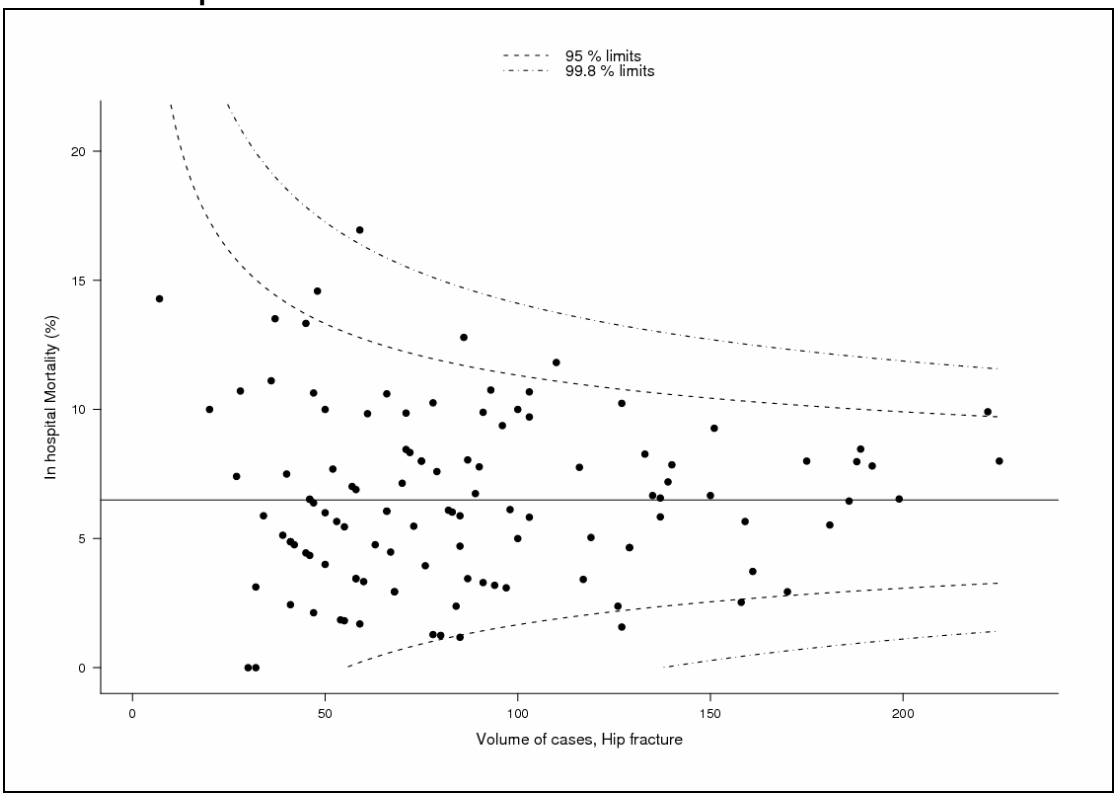
7.3.6 Volume outcome relationship

7.3.6.1 Analysis by hospital

IN-HOSPITAL MORTALITY, BY HOSPITAL

Figure 7.17 shows that only one hospital was outside the expected 99.8% limits of variability. The horizontal line represents the overall in-hospital mortality rate over all hospitals i.e. 6.5%.

Figure 7.17: Hip fracture surgery: Funnel plot of in-hospital mortality, by hospital



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE HOSPITAL

Table 7.42 shows the case mix and patient characteristics in function of the volume operated by hospital, classified by tertile. Regardless of the tertile, 75% of the patient population was female and had a mean age of 79.2 years. Compared with medium- and high-volume hospitals (more than 110 hip fracture surgeries in 2004), patients in small-volume hospitals (less than 60 procedures) had slightly different characteristics: there were more men (29.0%), they were slightly younger (mean age 78.1 years) and they had a higher severity index (34.1% with SOI = 3, 7.6% with SOI = 4). Their Charlson index score was similar, though.

Table 7.42: Hip fracture surgery: Differences in case-mix and outcomes by volume category of hospital

			Tertile based on volume			
			≤59	>59 and ≤96	>110	All
Number hospitals			39	37	37	113
Number stays			1712	2891	5331	9934
Male		%	29.0	25.3	24.0	25.3
Age		Mean	78.1	79.1	79.6	79.2
		Std	13.0	12.2	11.7	12.1
Charlson index score		Mean	1.2	1.3	1.2	1.2
		Std	1.6	1.7	1.7	1.7
Charlson ≥ 3		%	18.5	19.0	17.1	17.9
Severity index (APR-DRG)	1	%	23.8	22.3	21.9	22.3
	2	%	34.6	39.6	39.1	38.5
	3	%	34.1	30.9	31.7	31.9
	4	%	7.6	7.2	7.3	7.3
In-hospital mortality		%	6.4	6.3	6.6	6.5

RELATION BETWEEN HOSPITAL VOLUME AND IN-HOSPITAL MORTALITY

The results of the logistic regression are presented in Table 7.43. There is no significant association between a 10% increase in hospital volume and the in-hospital mortality as the effect was +0.48% (-1.22%, 2.17%), without adjusting for case-mix differences. After adjustment for case mix, an increase of 1 category Charlson score was translated by a more than doubling of the in-hospital mortality; the odds ratio being 2.13 (1.99, 2.28). Being a male gave the same effect (odds for mortality: 1.99; 1.64, 2.40). Both effects were statistically significant. On the contrary, an increase of 10% of the hospital volume still had no significant effect (+0.61%; -1.24%, 2.47%).

Table 7.43: Hip fracture surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of in-hospital mortality

Model without adjustment for case mix	Effect[1]	95% CI	
Hospital volume (increase of 10%)	0.48	-1.22	2.17
Model with adjustment for case mix	Effect[1]	95% CI	
Hospital volume (increase of 10%)	0.61	-1.24	2.47
	Odds Ratio	95% CI	
Sex (male vs. female)	1.99	1.64	2.40
Age (increase of 1 year)	1.06	1.05	1.07
Charlson score (increase of 1 category)	2.13	1.99	2.28

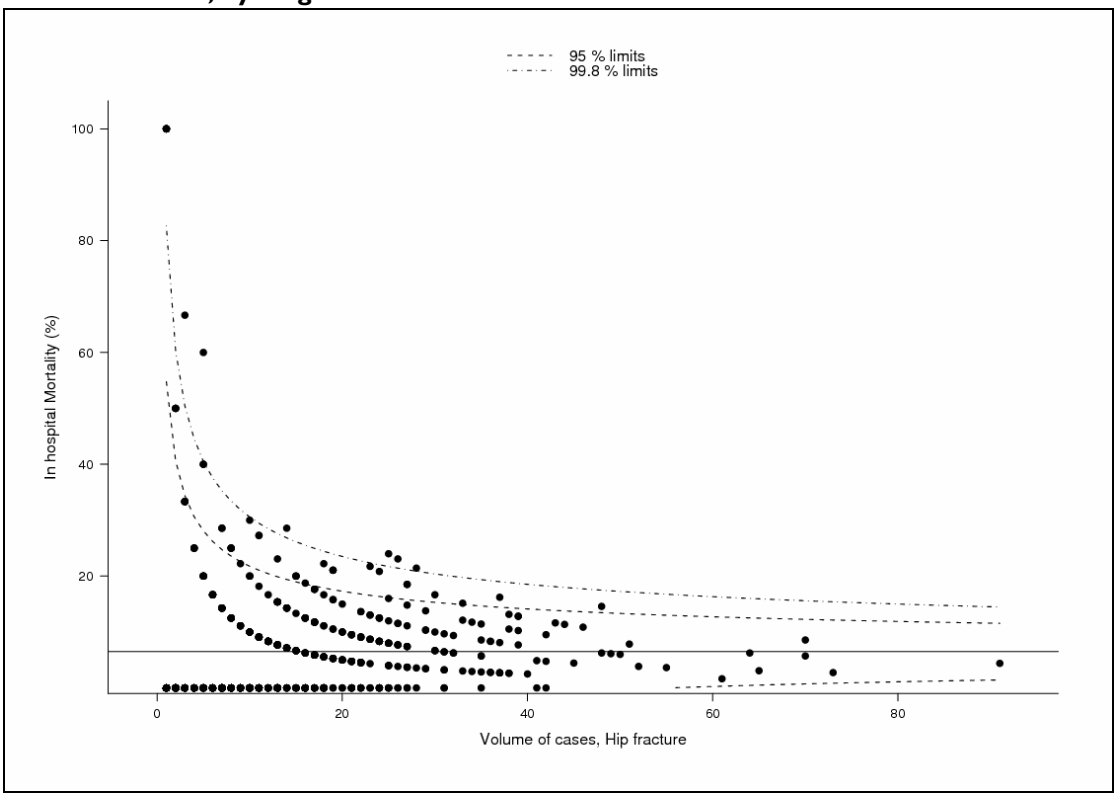
[1] Effect of 10% increase in volume of hospital on the odds of in-hospital mortality

7.3.6.2 *Analysis by surgeon*

IN-HOSPITAL MORTALITY, BY SURGEON

The funnel plot (Figure 7.18) of the in-hospital mortality in function of the number of operations done by surgeon shows that six relatively small to medium-sized hospitals were situated outside the limits of the expected variability.

Figure 7.18: Hip fracture surgery: Funnel plot of the in-hospital mortality rate, by surgeon



DESCRIPTION OF DIFFERENCES IN CHARACTERISTICS ACCORDING TO THE VOLUME CATEGORY OF THE SURGEON

Table 7.44 presents the patient characteristics by surgeon volume tertile. Patients operated by low-volume surgeons (i.e. less than 9 hip fracture surgeries in 2004) are slightly younger (mean age 77.8 years) and are more likely to be male (27.3%) than those operated by medium- or high-volume surgeons (more than 17 procedures per year). On the other hand, however, patients of low-volume surgeons have a lower Charlson index score and a lower severity of index. In-hospital mortality was very comparable between low- and high-volume surgeons i.e., respectively, 6.8 and 6.7%. Medium-volume surgeons, on the other hand, had slightly lower in-hospital mortality (6.0%).

Table 7.44: Hip fracture surgery: Differences in case-mix and outcomes by volume category of surgeon

			Tertile based on volume			
			≤8	> 8 and ≤17	> 17	All
Number surgeons			234	225	216	675
Number stays			960	2874	6100	9934
Male		%	27.3	24.6	25.2	25.3
Age		Mean	77.8	78.8	79.5	79.2
		Std	13.4	12.5	11.7	12.1
Charlson index score		Mean	1.1	1.2	1.3	1.2
		Std	1.6	1.6	1.7	1.7
Charlson ≥ 3		%	15.7	17.2	18.6	17.9
Severity index (APR-DRG)	1	%	23.6	23.1	21.8	22.3
	2	%	38.3	36.7	39.3	38.5
	3	%	30.7	32.6	31.7	31.9
	4	%	7.3	7.6	7.2	7.3
In-hospital mortality		%	6.8	6.0	6.7	6.5

RELATION BETWEEN SURGEON VOLUME AND IN-HOSPITAL MORTALITY

The effects from the logistic regression are presented in Table 7.45 for in-hospital mortality. Without adjustment for case mix, an increase of 10% of the surgeon had no significant effect on the in-hospital mortality as the effect was -0.05% (-1.33%, +1.23%). After adjustment, the effect is increased but is still not significant at -0.68% (-2.05%, 0.69%). The odds for mortality are doubled in case of male patient, they are slightly higher when age increased by 1 year and they are more than doubled by an increase of 1 category in the Charlson index score.

Table 7.45: Hip fracture surgery: Correlation-corrected logistic regression (GEE) estimates of determinants of in-hospital mortality (surgeon volume)

Model without adjustment for case mix	Effect[1]	95% CI	
Surgeon volume (increase of 10%)	-0.05	-1.33	1.23
Model with adjustment for case mix	Effect1	95% CI	
Surgeon volume (increase of 10%)	-0.68	-2.05	0.69
	Odds Ratio	95% CI	
Sex (male vs. female)	1.98	1.65	2.36
Age (increase of 1 year)	1.06	1.05	1.07
Charlson score (increase of 1 category)	2.13	1.98	2.29

[1] Effect of 10% increase in volume of surgeon on the odds of in-hospital mortality

7.3.7 Discussion

7.3.7.1 External validation of the definition of the procedure

According to the Dutch guidelines on osteoporosis from 2002, there would have been 15 290 patients above 55 years for hip fracture hospitalized in the Netherlands in 1999.[339] Considering the number of persons of more than 55 years, if the same incidence was observed in Belgium, there would be 12 120 hospitalizations of patients of more than 55 years, which is of course inferior to the 15 230 patients hospitalized for hip fracture identified by the data extraction, all ages included.

7.3.7.2 *Summarized results of literature review*

The systematic literature search identified three systematic reviews in which the VOA for hip fracture surgery was studied.[1, 5, 64] The systematic reviews identified 8 primary studies of which 4 were published between 1988 and 2000.[103, 309, 338, 340] These articles were complemented with a limited number of primary studies that were retrieved through an additional search of literature.[341, 342]

On the basis of these three systematic reviews, it was concluded in Chapter 2 (see Table 2.2 on page 19) that there was an inverse relation between hospital volume and mortality. Nevertheless, it has to be emphasised again that each systematic review had studied only a very limited number of primary studies.

Detailed results of the literature review, using the data extraction template, are described in the evidence tables in Appendix 10.

7.3.7.3 *Comparative analysis of literature and Belgian data*

DEFINITION OF VOLUME

Table 7.39 shows that Belgian hospitals performed on average 88 hip fracture surgeries, in 2004. Belgian surgeons had an annual mean of 15 hip fracture surgeries in 2004 (see Table 7.40). Wenning et al. mention a mean annual volume of 145 hip fracture surgeries in hospitals in North Rhine-Westphalia, Germany.[340] They studied data of 27 000 patients who needed hip fracture surgery between 1993 and 1998.

OUTCOME

In-hospital mortality after hip fracture surgery was 6.49% in this KCE study (see Table 7.41). As shown in Table 7.46, this is very similar to the 6.2% in-hospital mortality reported by Wenning.[340] Holt and colleagues analyzed data from the Scottish Hip Fracture Audit on more than 25 000 patients over 50 years old who are admitted to hospital with a hip fracture between 1998 and 2005.[342] They reported a mortality of 8.10% 30 days after surgery and of 20.2% at 120 days. Shah et al. analyzed more than 173 000 cases of hemi-arthroplasty for femoral neck fracture in patients ≥65 years of age.[341] They reported a significantly lower in-hospital mortality rate of 3.1%.

Table 7.46: Hip fracture surgery: in-hospital mortality: comparison of Volume-Outcome KCE study with scientific literature

KCE study Volume outcome	Wenning 2000	Holt 2008	Shah 2005
6.49%	6.20%	8.10% at 30 days	3.1%

PATIENT CASE MIX

Our results showed that patients operated in low-volume hospitals (less than 60 hip fracture surgeries in 2004) or by low-volume surgeons (less than 9 hip fracture procedures), are more likely to be male and are slightly younger (see Table 7.42 and Table 7.44). In low-volume hospitals, patients seemed to have a higher severity of illness index, although the Charlson index scores were similar compared with the other volume tertiles. Patients operated by low-volume surgeons, on the other hand, had a lower Charlson index score and a lower severity of index.

There are very few studies to compare with. Hamilton et al. analysed longitudinal data from all hospitals in Quebec that carried out hip fracture surgery between 1990 and 1993.[338] They found that high-volume hospitals (≥73 procedures per year) had patients with significantly fewer co-morbidities and that these co-morbidities were less severe (lower Charlson score).

VOLUME OUTCOME ASSOCIATION

Our results show that, after case-mix adjustment, hospitals with a higher volume of hip fracture surgeries are not associated with lower in-hospital mortality (see Table 7.43). The same accounts for surgeon volume (see Table 7.45).

Other studies did find an inverse relation between higher volumes and outcomes. Nevertheless, the studies by Hughes and by Taylor were not suitable for comparison.[103, 309] First, their patient population was not analogous to ours: Hughes included hip fracture patients that had no surgery as well as those with fractures of other (than the neck) or unspecified parts of the femur; Taylor studied all patients who underwent a hip procedure whether they had a hip fracture or not. Second, these studies did not adjust for case-mix.

Wenning et al. found that patients in low-volume hospitals (less than 15 hip procedures per year) were more at risk of dying postoperatively (odds ratio is 1.33; 95% CI 1.09, 1.63) compared to those in high-volume hospitals (more than 45 procedures per year).[340] Shah et al. limited their study to patients ≥65 years of age with femoral neck fracture undergoing hemi-arthroplasty (ICD-9-CM procedure code 81.52), in the USA between 1988 and 2000.[341] He demonstrated that patients of low-volume surgeons (1 to 3 procedures per year) had an 18% (odds ratio 1.18; 95% CI 1.03, 1.34) increased risk of mortality compared with patients treated by surgeons performing ≥12 procedures per year. Hospital volume, on the other hand, could not predict mortality. Hamilton and colleagues, on the other hand, showed that fluctuations in a hospital's volume from period to period had no significant effect on mortality.[338]

Although we did not find a volume outcome association, the logistic regressions showed that male patients are almost two times more likely to die during the index admission than female patients (odds ratio is 1.99 in Table 7.43 and 1.98 in Table 7.45). Holt et al. reported similar gender differences in epidemiology and outcome after hip fracture.[342] Twenty-two percent of these patients was male, while 78% was female. Men had a younger mean age at presentation (i.e. 77 years compared with 81 years for women). Despite the fact that male patients were younger, they were in poorer pre-operative health (i.e. men were more likely to be ASA 3 or above). Mortality at 30 and 120 days was almost two times higher for men than for women, even after controlling for the effects of case-mix variables.

Key points on volume outcome association for hip fracture surgery (HFS)

- **A total of 9 934 patients were hospitalized in 2004 for a fracture of the neck of the femur and underwent a surgical intervention.**
- **These interventions were performed in 113 centres and by 675 surgeons. Annual mean number of HFS was 88 per centre, and 15 per surgeon. These volumes are much higher than US data (44% of centres less than 11 elective HFS).**
- **In-hospital mortality is 6.5% after HFS. Logistic regression models were fitted to assess association between this outcome and centre or surgeon volume. Factors included in the models were: age, sex, Charlson score, and principal diagnosis (ostheoarthrosis versus others diagnoses).**
- **Although the literature study concluded that there was an inverse relation between hospital volume and mortality, in Belgian data neither hospital or surgeon volume were associated with in-hospital mortality after HFS.**
- **Unfortunately, the MCD do not provide information on the outcomes of greatest interest to patients such as loss of independence, loss of mobility or residual pain.**

8 DISCUSSION

Why a report on the volume outcome association in Belgium? This topic has been studied in the US since the start of the 80s, and in other countries afterwards. Despite suffering sometimes of some methodological shortcomings, many of these studies have shown that, for specific procedures, patients admitted to low-volume hospitals or treated by low-volume surgeons have worse outcomes than patients admitted to high-volume hospitals or treated by high-volume surgeons. In Belgium, very few studies have been performed on that subject, probably due to the difficult access to the required data.

This study is thus a pioneer in our country, and has been designed first of all as a feasibility study. The main research question was: Is it possible to use administrative hospital data to examine the volume outcome relationship? To answer that question, we selected 13 interventions from three medical domains, where literature on the volume outcome association was abundant: oncology, cardiology and orthopaedic surgery. This wide variety of domains and interventions allowed us to draw global conclusions on the use of those volume indicators in the framework of improvement of quality of care.

The answer to the above mentioned objective is cautiously positive: this study shows that Belgian administrative hospital data can be used to study the volume outcome relationship provided all available information is retrieved from the databases, and, preferably, linked to clinical registries. Linkage with clinical data would be only one way to improve the risk adjustment.

The main **limitation of administrative data** is the possibility to define the **outcome of interest** for each procedure. Many procedures that were analyzed in the literature could not be analyzed on our data because of the lack of information on the outcome. Examples include incontinence and quality of life after transurethral prostatectomy, adhesions in women who had a caesarean section, loss of mobility and residual pain after total hip arthroplasty.

For those procedures where it is possible to define the outcome of interest, the necessary precautions that should be taken in the analysis are listed in the 9 points below.

1. Great care is needed in **identifying the study population** in the administrative databases. Surgical procedures are coded with two different coding systems: the ICD-9 classification in the minimal clinical data (MCD) and the NIHDI billing codes (nomenclature) in the minimal financial data (MFD). There is no 1:1 equivalence between these two coding systems, and, depending on the intervention, the ICD-9, the nomenclature, or both are needed to obtain a precise description of the procedure. The reason for intervention (i.e. principal diagnosis in MCD) is also necessary to include or exclude specific groups of patients.
2. Serious thoughts must be given to the **time horizon**, i.e. the time between the intervention and the evaluation of the outcome. The choice of time horizon depends on whether the intervention is complex and therefore high risk, and whether the interest is on the surgeon volume or the centre volume.
    a. For complex and therefore high risk procedures such as oesophageal cancer surgery or CABG, outcomes can be assessed at short term (in-hospital, 3-months or 6–months mortality). For less complex procedures or conditions with a good prognosis, such as breast cancer, outcome cannot be assessed at short term simply because there are not enough events. In these cases, evaluation has to be performed in a longer-term perspective, keeping in mind that other treatments besides surgery affect patient's survival. Ideally, the outcome measure should be adapted for each procedure.

b. Evaluation of surgeon volume can be performed at short term: in-hospital mortality, 30-day or 90-day mortality, keeping in mind that evaluating outcome beyond hospital discharge has the additional advantage of being unaffected by differences in practice discharges across hospitals. Especially for low-risk surgical procedures, evaluation of centre volume can require several years of follow up, because the effect of centre volume is a mixture of the effect of the surgeon volume (the experience) and the effect of organisational aspects of the centres (such as process indicators, compliance to guidelines, organization of care), which play a role on a longer timeframe than surgery skills. To analyze outcomes beyond hospitalisation, it is necessary to link MCD-MFD data to IMA data. One shortcoming of these data is that the exact dates (of hospital admission and of death) are not available, hence rending impossible to evaluate outcomes at precise time points.

3. It is important to **distinguish between the effect of the surgeon volume (experience) and the effect of the hospital volume (organisation of care in the broad sense)**. The relative importance of surgeon or hospital volume is difficult to distinguish for infrequent interventions where surgeon volume equals hospital volume. In addition, this relative importance seems to vary according to the procedure. Two extremes are carotid endarterectomy and lung cancer surgery. CEA, for example, is technically demanding and any failure in surgical practice is potentially catastrophic. Other hospital-based services, on the other hand, are relatively less important i.e. most patients undergoing CEA do not require intensive postoperative management. In the case of lung cancer surgery, in contrast, patients rarely die because of direct technical complications of the procedure itself. Since these patients more often die from cardiac events, pneumonia and respiratory failure, hospital-based services are very important. These services include, for example, intensive care, pain management, respiratory care and nursing care.

4. **Robust information on case mix** is important. Patient characteristics and disease severity should be available. Risk adjustment is an important issue in volume outcome research because patients with severe co morbidity may be unequally distributed between providers of low and high volume.

   a. MCD data provide information on patient demographics and co morbidities. A useful tool is the Charlson score, which has been validated to predict 1-year mortality. It is a sum of some predefined weights attributed to 17 specific conditions. The Charlson score can be computed based on MCD data, but inherits their limitations i.e. it depends of the quality and completeness of coding of co morbidities in each hospital. The other drawback of MCD data, the impossibility to differentiate co morbidities from complications, has been resolved in MCD since 2008.

   b. The other part of the case mix adjustment, information on disease severity, is not available in MCD data but can be retrieved in existing registries, which typically record detailed clinical information.

   - With respect to cancer, the Belgian Cancer Registry has detailed information on tumour characteristics, and the linkage in our study between MCD and BCR data for five conditions was successful. Stage was still missing in a high proportion of cases (on data of 2004), but efforts are currently made to enhance completeness of stage coding. In future, the availability of an updated and complete register in combination with the patient's personal identification number will enable complete follow-up of cancer patients in Belgium.

- In cardiology, the Belgian Working Group for Interventional Cardiology (BWGIC) disposes of data for CABG and PCI, and, although we did not link MCD data to BWGIC registry due to time constraints, this type of linkage has already been performed in the past and could be redone.
- For orthopaedic procedures, only very recent the National Institute for Health and Disability Insurance (NIHDI) started the electronic registry ORTHOpedic (Prosthesis Identification Data ORTHOpride) which will include all prostheses of hip and knee.[343] Such initiatives will definitely be of benefit to the quality of future studies.

We cannot rule out confounding by unmeasured characteristics of patients in our study. Nevertheless, we do not believe that limitations related to risk adjustment threaten our main conclusions about the association between volume and outcome.

5. **Appropriate statistical methods** should be used. Regression models are available that respect the hierarchical nature of the data (patients nested within surgeons, surgeons nested within centres), and account for the correlations within these clusters. The failure to include any type of adjustment for those correlations would lead to falsely high statistically significant effects. Generalized Estimating Equations (GEE) and Generalized Linear Mixed Models (GLMM) and are two examples of methods that can be used to analyze hierarchical data.

6. The graphical presentation of results. The **funnel plot** is a good and "easy to use" tool. It avoids spurious ranking of institutions, spurious stigmatization of low volume centres, and allows for an informal assessment of any volume outcome relationship.

7. **Sensitivity analyses and robustness of the results**. Results should be transparent. In our study, effects of volume were always presented with and without adjustment for case mix (based on administrative data only or based on all clinical data available) so that one can assess the influence of case mix on the volume effect. Sensitivity analyses were also performed when there were huge volume outliers, where it is difficult to differentiate the effect of the volume or other characteristics of that centre.

8. Problem of **missing data on cancer stage**: on average 30% missing data, sometimes 40% in small volume hospitals. The fact that mortality rate was not substantially higher in the patients whose stage was missing in the BCR seems to indicate, that, in this study, these patients are randomly divided into the four disease stages. Ideally, though, this assumption should have been checked with the help of sensitivity analyses, but this was not done due to time constraints.
   In addition, we noticed that many hospitals – low-volume as well as high-volume – missed data on stage and that the percentage of missing data varied among these hospitals. Despite the failure to retrieve information on disease stage, this problem did not restrain us from drawing conclusions on the volume outcome association. Nevertheless, this problem of missing data on disease stage (and other variables useful for risk adjustment) supports the need for complete and accurate data collection.

9. **Sample size** is sometimes not sufficient in one year: analysis of several years for rare tumours or procedures (pancreas, oesophagus, heart transplant) is required.

When all those precautions are taken into consideration and when associations between volume and outcome are demonstrated, this does not guarantee that there is a causal relationship between the two. A competing hypothesis is the "selective referral hypothesis", whereby good quality hospitals have also a high volume because they attract more patients who are aware of their quality. Unlike in US or in UK, outcome data are not publicly available in Belgium, which renders the mechanism behind the "selective referral hypothesis" difficult to support. There are also statistical methods to disentangle the two hypotheses, but they rely themselves heavily on other assumptions that can hardly be verified.

What is the link between outcome and quality of care? Outcome indicators are only one component of quality, process indicators are also needed. This study was limited to outcome indicators, but processes of care could also be analyzed based on administrative data. In the case of cancer, the administration of adjuvant therapy is documented in the IMA-AIM database and could be used to analyze if appropriate use of adjuvant therapy is indeed better in higher-volume hospitals which could help explain the volume outcome association. It has to be acknowledged though that certain aspects of adjuvant therapy are not documented in the IMA data i.e. exact time of administration of the drug and information on whether the course of therapy was completed.

In other countries, proofs of relationship between high volume of interventions and good outcome have led to centralisation of care. In addition, in a limited number of countries, minimal volume-based criteria (between brackets) were issued for a number of operations. In the USA, the Leapfrog group already has annual volume thresholds for CABG (450), PCI (400), abdominal aortic aneurysm repair (50), aortic valve replacement (120), pancreatic resection (11), oesophagectomy (13) and bariatric surgery (125).[126] The Leapfrog Group is a non-profit organisation which represents many of the US' largest private and public-sector healthcare purchasers. The coalition encourages both patients and payers to select hospitals that meet minimal volume standards for the abovementioned high-risk procedures. The American Agency for Healthcare Research and Quality (AHRQ) defined similar volume thresholds for 6 procedures i.e. oesophageal resection, pancreatic resection, abdominal aortic aneurysm repair, CABG, PCI and CEA.[29, 248]

It is reassuring, though, to see that these volume cut offs are under constant revision and are updated when necessary. CEA, for example, has recently been removed from the Leapfrog list due to updated evidence that showed only moderate effect of volume. Other thresholds have been revised downwards: the Leapfrog cut off of 450 CABG per year is currently criticized and a smaller cut off of 200 is proposed.[33, 126]

In the UK, the evidence of improved outcome in high-volume hospitals led to governmental guidance in the form of the Department of Health's 2001 document: Improving outcomes in upper gastrointestinal cancers.[344] Implementation of the recommendations of this document generated a major change in the delivery of gastrointestinal surgery: oesophago-gastric and hepato-pancreato-biliary surgery were centralised into centres with large catchment populations.[345]

But there are objections to such initiatives aimed at concentrating certain surgical procedures in high-volume hospitals. It may be pointed out that procedure volume is an imperfect proxy for quality and that some low-volume hospitals have excellent outcomes, while some high-volume hospitals have poor outcomes. Other barriers to implementation of a selective referral programme include patients' preferences for care near home, inability to transfer unstable patients to high-volume centres etc. Many of these aspects were discussed at a workshop organised by the Institute of Medicine in the year 2000.[346] The total number of patients involved also plays a crucial role: centralisation of a rare procedure (a few hundred patients per year) completely differs from centralisation of a very common procedure (more than 10 000 patients per year). Centralisation has consequences on the organisation of health care, and those consequences should not be neglected.

Regardless of the desirability of centralisation, it is clear that no programme will ever succeed in getting all high-risk procedures into high-volume centres. Therefore, some authors argue that instead of centralisation, it is probably worthwhile to develop strategies for improving care in ALL hospitals, even low volume ones. This should be done by setting a goal of improving surgical quality of care by improving the processes of care.

Anyway, the merits and drawbacks of centralisation of health care are beyond the scope of this report. Our analyses were not intended to determine thresholds but were primarily designed to assess the feasibility of a volume outcome study with Belgian administrative data and to validate the existence of volume-outcome associations.

The fact that this study is based on relatively old data (2004) calls for some caution in the effects observed. Many changes have been introduced in our health care system since 2004. Fusion of hospitals, creation of Belgian cancer registry, introduction of the multidisciplinary consultations in oncology, introduction of minimal volumes in the treatment of breast cancer, etc… All those factors have a potential impact on the quality of care. All our results are thus valid for the year 2004, but might not be applicable now.

Our final conclusion is that Belgian administrative data linked to registry data can be an adequate tool to study the volume outcome relationship, provided necessary precautions are taken.

9 REFERENCES

1. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med. 2002;137(6):511-20.
2. Van Den Bruel A, Vlayen J, Gailly J, Chalon P, Ramaekers D. Good Clinical Practice (GCP). Process notes. Brussels: Belgian Health Care Knowledge Centre (KCE); 2007. KCE Process notes (D2007/10.273/38)
3. NHS Centre for Reviews and Dissemination. Relationship between volume and quality of health care: a review of the literature. York: University of York; 1995. CRD Report 2 Available from: http://www.crd.york.ac.uk/CRDWeb/ShowRecord.asp?View=Full&ID=31999008337
4. Hannan EL. Evaluating and improving the quality of care for acute myocardial infarction: can regionalization help? Jama. 2006;295(18):2177-9.
5. Gandjour A, Bannenberg A, Lauterbach KW. Threshold volumes associated with higher survival in health care: a systematic review. Med Care. 2003;41(10):1129-41.
6. Guyatt G, Gutterman D, Baumann MH, Addrizzo-Harris D, Hylek EM, Phillips B, et al. Grading strength of recommendations and quality of evidence in clinical guidelines: report from an american college of chest physicians task force. Chest. 2006;129(1):174-81.
7. Biewener A, Aschenbrenner U, Sauerland S, Zwipp H, Rammelt S, Sturm J. Impact of rescue pathway and hospital volume on the mortality of polytraumatized patients. Status report. Unfallchirurg. 2005;108(5):370-7.
8. Urbach DR, Croxford R, MacCallum NL, Stukel TA. How are volume-outcome associations related to models of health care funding and delivery? A comparison of the United States and Canada. World J Surg. 2005;29(10):1230-3.
9. Banta D, Bos M. The relation between quantity and quality with coronary artery bypass graft (CABG) surgery. Health Policy. 1991;18(1):1-10.
10. Banta HD, Engel GL, Schersten T. Volume and outcome of organ transplantation. Int J Technol Assess Health Care. 1992;8(3):490-505.
11. Bentrem DJ, Brennan MF. Outcomes in oncologic surgery: does volume make a difference? World J Surg. 2005;29(10):1210-6.
12. Black N, Johnston A. Volume and outcome in hospital care: evidence, explanations and implications. Health Serv Manage Res. 1990;3(2):108-14.
13. Black PC, Brown GA, Dinney CP. Should cystectomy only be performed at high-volume hospitals by high-volume surgeons? Curr Opin Urol. 2006;16(5):344-9.
14. Casson AG, van Lanschot JJ. Improving outcomes after esophagectomy: the impact of operative volume. J Surg Oncol. 2005;92(3):262-6.
15. Chang AC, Birkmeyer JD. The volume-performance relationship in esophagectomy. Thorac Surg Clin. 2006;16(1):87-94.
16. Dibra A, Kastrati A, Schuhlen H, Schomig A. The relationship between hospital or operator volume and outcomes of coronary patients undergoing percutaneous coronary interventions. Z Kardiol. 2005;94(4):231-8.
17. Dimick JB, Birkmeyer JD, Upchurch GR, Jr. Measuring surgical quality: what's the role of provider volume? World J Surg. 2005;29(10):1217-21.
18. Keck T, Makowiec F, Adam U, Hopt UT. [Does hospital volume have influence on the results of pancreatic surgery?]. Zentralbl Chir. 2007;132(1):26-31.
19. Kelly JV. What do mortality studies reveal about hospital volume, teaching status, and ownership? Int J Technol Assess Health Care. 1990;6(2):239-52.
20. Nathens AB, Maier RV. The relationship between trauma center volume and outcome. Adv Surg. 2001;35:61-75.
21. Nguyen NT, Higa K, Wilson SE. Improving the quality of care in bariatric surgery: the volume and outcome relationship. Adv Surg. 2005;39:181-91.
22. Snider RL, Laskey WK. Quality management and volume-related outcomes in the cardiac catheterization laboratory. Cardiol Clin. 2006;24(2):287-97, vii.
23. Bennett CL, Deneffe D. Does experience improve hospital performance in treating patients with AIDS? Health Policy. 1993;24(1):35-43.

24. Hutter MM. Does outcomes research impact quality? Examples from bariatric surgery. Am Surg. 2006;72(11):1055-60; discussion 61-9, 133-48.
25. Watson RS. Location, location, location: regionalization and outcome in pediatric critical care. Curr Opin Crit Care. 2002;8(4):344-8.
26. Steinbruchel DA, Ravn JB. [The connection between hospital volume and outcome in thorax surgery]. Ugeskr Laeger. 2006;168(15):1524-6.
27. Schlenker RE, Hittle DF, Hrincevich CA, Kaehny MM. Volume/outcome relationships in small rural hospitals. J Rural Health. 1996;12(5):395-409.
28. Khuri SF, Henderson WG. The case against volume as a measure of quality of surgical care. World J Surg. 2005;29(10):1222-9.
29. Agency for Healthcare Research and Quality. Guide to Inpatient Quality Indicators: Quality of Care in Hospitals - Volume, Mortality, and Utilization. Rockville, MD: Agency for Healthcare Research and Quality (AHRQ); 2007. AHRQ Quality Indicators Available from: http://www.qualityindicators.ahrq.gov/downloads/iqi/iqi_guide_v31.pdf
30. Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ. The volume-outcome relationship: don't believe everything you see. World J Surg. 2005;29(10):1241-4.
31. Finlayson SR. The volume-outcome debate revisited. Am Surg. 2006;72(11):1038-42; discussion 61-9, 133-48.
32. Kraus TW, Buchler MW, Herfarth C. Relationships between volume, efficiency, and quality in surgery--a delicate balance from managerial perspectives. World J Surg. 2005;29(10):1234-40.
33. Shahian DM, Normand SL. The volume-outcome relationship: from Luft to Leapfrog. Ann Thorac Surg. 2003;75(3):1048-58.
34. Rogers SO, Jr. The holy Grail of surgical quality Improvement: process measures or risk-adjusted outcomes? Am Surg. 2006;72(11):1046-50; discussion 61-9, 133-48.
35. Davoli M, Amato L, Minozzi S, Bargagli AM, Vecchi S, Perucci CA. [Volume and health outcomes: an overview of systematic reviews]. Epidemiol Prev. 2005;29(3-4 Suppl):3-63.
36. Burgers JS, Wittenberg J, Kallewaard M, van Croonenborg JJ, van Barneveld TA, van Everdingen JJ. [Relationship between volume and quality of care for surgical interventions; results of a literature review]. Ned Tijdschr Geneeskd. 2007;151(38):2105-10.
37. Gandjour A, Lauterbach KW. [The relationship between provider volume and mortality rate: volume data of German centres of excellence]. Z Kardiol. 2001;90(9):613-20.
38. Halm EA, Lee C, Chassin MR. How is volume related to quality in health care? A systematic review and methodologic critique of the literature. In: Hewitt M, editor. Interpreting the Volume-Outcome Relationship in the Context of Health Care Quality: Workshop Summary. Washington DC: Institute of Medicine; 2000. p. S27-102. Available from: http://www.nap.edu/catalog.php?record_id=10005
39. Hillner BE. Is cancer care best at high-volume providers? Curr Oncol Rep. 2001;3(5):404-9.
40. NHS Centre for Reviews and Dissemination. Hospital volume and health care outcomes, costs and patient access. York: University of York; 1996. Effective Health Care 2(8) Available from: http://www.york.ac.uk/inst/crd/ehc28.pdf
41. NHS Centre for Reviews and Dissemination. Concentration and Choice in the Provision of Hospital Services. Summary report. York: University of York; 1997. CRD Report 8 Available from: http://www.crd.york.ac.uk/CRDWeb/ShowRecord.asp?ID=31999008334
42. Sowden A, Aletras V, Place M, Rice N, Eastwood A, Grilli R, et al. Volume of clinical activity in hospitals and healthcare outcomes, costs, and patient access. Qual Health Care. 1997;6(2):109-14.
43. Sowden AJ, Deeks JJ, Sheldon TA. Volume and outcome in coronary artery bypass graft surgery: True association or artefact? British Medical Journal. 1995;311(6998):151-5.
44. Teisberg P, Hansen FH, Hotvedt R, Ingebrigtsen T, Kvalvik AG, Lund E, et al. Pasientvolum og behandlingskvalitet. Oslo: Norwegian Knowledge Centre for the Health Services (NOKC); 2001. SMM-Report 2/2001 Available from: http://www.kunnskapssenteret.no/filer/Rapport2-01.pdf
45. Pla R, Pons JM, González JR, Borràs JM. ¿Hay niveles asistenciales en cirugía oncológica? ¿Los que más hacen, lo hacen mejor? Informe sobre la relación entre volumen de procedimientos y resultados en cirugía oncológica. Barcelona: Agència d'Avaluació de Tecnologia i Recerca

Mèdiques. Catsalut. Departament de Sanitat i Seguretat Social. Generalitat de Cataluña; 2003. Available from: http://www.gencat.net/salut/depsan/units/aatrm/html/en/dir393/doc7921.html

46. Battaglia TC, Mulhall KJ, Brown TE, Saleh KJ. Increased surgical volume is associated with lower THA dislocation rates. Clin Orthop Relat Res. 2006;447:28-33.
47. Brisman JL, Niimi Y, Song JK, Berenstein A. Aneurysmal rupture during coiling: low incidence and good outcomes at a single large volume center. Neurosurgery. 2005;57(6):1103-9; discussion -9.
48. Hebert-Croteau N, Roberge D, Brisson J. Provider's volume and quality of breast cancer detection and treatment. Breast Cancer Res Treat. 2007;105(2):117-32.
49. Hillner BE, Smith TJ, Desch CE. Hospital and physician volume or specialization and outcomes in cancer treatment: importance in quality of cancer care. J Clin Oncol. 2000;18(11):2327-40.
50. Holscher AH, Metzger R, Brabender J, Vallbohmer D, Bollschweiler E. High-volume centers--effect of case load on outcome in cancer surgery. Onkologie. 2004;27(4):412-6.
51. Le Pimpec-Barthes F, Bagan P, Hubsch JP, Bry X, Pereira Das Neves JC, Riquet M. [Evaluation of thoracic surgical practice. The impact of specialisation and the effect of volume on the results of cancer treatment: resectability, post-operative mortality, and long-term survival]. Rev Mal Respir. 2006;23(4 Suppl):13S73-85; quiz 13S157, 13S9.
52. Meyer HJ. The influence of case load and the extent of resection on the quality of treatment outcome in gastric cancer. Eur J Surg Oncol. 2005;31(6):595-604.
53. Obladen M. [Minimum patient volume in care for very low birthweight infants: a review of the literature]. Z Geburtshilfe Neonatol. 2007;211(3):110-7.
54. Slim K, Flamein R, Chipponi J. [Relation between activity volume and surgeon's results: myth or reality?]. Ann Chir. 2002;127(7):502-11.
55. Weitz J, Koch M, Friess H, Buchler MW. Impact of volume and specialization for cancer surgery. Dig Surg. 2004;21(4):253-61.
56. Shervin N, Rubash HE, Katz JN. Orthopaedic procedure volume and patient outcomes: a systematic literature review. Clin Orthop Relat Res. 2007;457:35-41.
57. Van Brabandt H, Camberlin C, Vrijens F, Parmentier Y, Ramaekers D, Bonneux L. Variaties in de ziekenhuispraktijk bij acuut myocardinfarct in België. Report. Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2005 23 juni. KCE reports 14A (D/2005/10.273/11) Available from: http://kce.fgov.be/Download.aspx?ID=265
58. Personal Communication by the Belgian Cancer Registry. In; 13/05/2009.
59. 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):145-61.
60. Dudley RA, Johansen KL, Brand R, Rennie DJ, Milstein A. Selective referral to high-volume hospitals: estimating potentially avoidable deaths. Jama. 2000;283(9):1159-66.
61. Holt PJ, Poloniecki JD, Loftus IM, Thompson MM. Meta-analysis and systematic review of the relationship between hospital volume and outcome following carotid endarterectomy. Eur J Vasc Endovasc Surg. 2007;33(6):645-51.
62. Khuri SF, Hussaini BE, Kumbhani DJ, Healey NA, Henderson WG. Does volume help predict outcome in surgical disease? Adv Surg. 2005;39:379-453.
63. Killeen SD, Andrews EJ, Redmond HP, Fulton GJ. Provider volume and outcomes for abdominal aortic aneurysm repair, carotid endarterectomy, and lower extremity revascularization procedures. J Vasc Surg. 2007;45(3):615-26.
64. NHS Centre for Reviews and Dissemination. Concentration and Choice in the Provision of Hospital Services. The relationship between hospital volume and quality of health outcomes. York: University of York; 1997. CRD Report 8 (part I) Available from: http://www.crd.york.ac.uk/CRDWeb/ShowRecord.asp?ID=31999008334
65. Shackley P, Slack R, Booth A, Michaels J. Is there a positive volume-outcome relationship in peripheral vascular surgery? Results of a systematic review. Eur J Vasc Endovasc Surg. 2000;20(4):326-35.
66. Kalant N, Shrier I. Volume and outcome of coronary artery bypass graft surgery: are more and less the same? Can J Cardiol. 2004;20(1):81-6.

67. IQWiG. Zusammenhang zwischen Menge der erbrachten Leistungen und der Ergebnisqualität für die „Percutane Transluminale Coronare Angioplastie (PTCA)“ Köln: 2006. Available from: http://www.iqwig.de/publikationen-des-iqwig.114.html
68. Stengel D, Ekkernkamp A, Dettori J, Hanson B, Sturmer KM, Siebert H. A rapid review of associations between provider volume and outcome of total knee arthroplasty. Where do the magical threshold values come from? Unfallchirurg. 2004;107(10):967-88.
69. 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.
70. Metzger R, Bollschweiler E, Vallbohmer D, Maish M, DeMeester TR, Holscher AH. High volume centers for esophagectomy: what is the number needed to achieve low postoperative mortality? Dis Esophagus. 2004;17(4):310-4.
71. Tjarda Van Heek N, Kuhlmann KFD, Scholten RJ, De Castro SMM, Busch ORC, Van Gulik TM, et al. Hospital volume and mortality after pancreatic resection: A systematic review and an evaluation of intervention in The Netherlands. Annals of Surgery. 2005;242(6):781-90.
72. Van Geenen RC, Gouma DJ. Impact of hospital volume on in-hospital mortality in pancreatic surgery. Surg Technol Int. 2002;10:61-5.
73. Hodgson DC, Fuchs CS, Ayanian JZ. Impact of patient and provider characteristics on the treatment and outcomes of colorectal cancer. J Natl Cancer Inst. 2001;93(7):501-15.
74. Iversen LH, Harling H, Laurberg S, Wille-Jorgensen P. Influence of caseload and surgical speciality on outcome following surgery for colorectal cancer: a review of evidence. Part 2: long-term outcome. Colorectal Dis. 2007;9(1):38-46.
75. Iversen LH, Harling H, Laurberg S, Wille-Jorgensen P. Influence of caseload and surgical speciality on outcome following surgery for colorectal cancer: a review of evidence. Part 1: short-term outcome. Colorectal Dis. 2007;9(1):28-37.
76. Henebiens M, van den Broek TA, Vahl AC, Koelemay MJ. Relation between hospital volume and outcome of elective surgery for abdominal aortic aneurysm: a systematic review. Eur J Vasc Endovasc Surg. 2007;33(3):285-92.
77. Holt PJ, Poloniecki JD, Gerrard D, Loftus IM, Thompson MM. Meta-analysis and systematic review of the relationship between volume and outcome in abdominal aortic aneurysm surgery. Br J Surg. 2007;94(4):395-403.
78. IQWiG. Zusammenhang zwischen der Menge der erbrachten Leistungen und der Ergebnisqualität für die Indikation „Elektiver Eingriff Bauchaortenaneurysma“. Köln: 2006 Mai 2006. Abschlussbericht Q05-01A Available from: http://www.iqwig.de/publikationen-des-iqwig.114.html
79. Nuttall M, van der Meulen J, Phillips N, Sharpin C, Gillatt D, McIntosh G, et al. A systematic review and critique of the literature relating hospital or surgeon volume to health outcomes for 3 urological cancer procedures. J Urol. 2004;172(6 Pt 1):2145-52.
80. Wilt T, Lederle F, MacDonald R, Jonk Y, Rector T, Kane R. Comparison of Endovascular and Open Surgical Repairs for Abdominal Aortic Aneurysm. Rockville: MD. Agency for Healthcare Research and Quality (AHRQ); 2006. Evidence Report/Tecnology Assessment No. 144
81. Wilt TJ, Shamliyan TA, Taylor BC, MacDonald R, Kane RL. Association between hospital and surgeon radical prostatectomy volume and patient outcomes: a systematic review. J Urol. 2008;180(3):820-8; discussion 8-9.
82. Luft HS, Garnick D, Mark D, McPhee S. Hospital volume, physician volume, and patient outcomes Perspectives HAP, editor.: Ann Arbor; 1990.
83. Seider H. Volume-outcome and Its Impact on U.S. Health Care Markets Beitraege G, editor.
84. Tsai AC, Votruba M, Bridges JF, Cebul RD. Overcoming bias in estimating the volume-outcome relationship. Health Serv Res. 2006;41(1):252-64.
85. Urbach DR, Baxter NN. Does it matter what a hospital is "high volume" for? Specificity of hospital volume-outcome associations for surgical procedures: analysis of administrative data. BMJ. 2004;328(7442):737-40.
86. Betensky RA, Christian CK, Gustafson ML, Daley J, Zinner MJ. Hospital volume versus outcome: an unusual example of bivariate association. Biometrics. 2006;62(2):598-604.
87. Adab P, Rouse AM, Mohammed MA, Marshall T. Performance league tables: the NHS deserves better. Bmj. 2002;324(7329):95-8.

88. Mohammed MA, Cheng KK, Rouse A, Marshall T. Bristol, Shipman, and clinical governance: Shewhart's forgotten lessons. Lancet. 2001;357(9254):463-7.

89. Spiegelhalter DJ. Funnel plots for comparing institutional performance. Stat Med. 2005;24(8):1185-202.

90. Flowers J, Hall P, Pencheon D. Public health indicators. Public Health. 2005;119(4):239-45.

91. Goldstein S. Statistical aspects of institutional performance: league tables and their limitations. Journal of Royal Statistical Society Series A. 1996(159):385-444.

92. Spiegelhalter DJ. Mortality and volume of cases in paediatric cardiac surgery: retrospective study based on routinely collected data. Bmj. 2002;324(7332):261-3.

93. Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ. The Leapfrog volume criteria may fall short in identifying high-quality surgical centers. Ann Surg. 2003;238(4):447-55; discussion 55-7.

94. Urbach DR, Austin PC. Conventional models overestimate the statistical significance of volume-outcome associations, compared with multilevel models. J Clin Epidemiol. 2005;58(4):391-400.

95. Iezzoni LI. Risk adjustment for measuring health care outcomes Press HA, editor. Chicago, Ill; 2003.

96. Panageas KS, Schrag D, Riedel E, Bach PB, Begg CB. The effect of clustering of outcomes on the association of procedure volume and surgical outcomes. Ann Intern Med. 2003;139(8):658-65.

97. Panageas KS, Schrag D, Russell Localio A, Venkatraman ES, Begg CB. Properties of analysis methods that account for clustering in volume-outcome studies when the primary predictor is cluster size. Stat Med. 2007;26(9):2017-35.

98. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. N Engl J Med. 1979;301(25):1364-9.

99. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? Health Serv Res. 1987;22(2):157-82.

100. McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. Jama. 1994;272(11):859-66.

101. Gowrisankaran G, Town RJ. Estimating the quality of care in hospitals using instrumental variables. J Health Econ. 1999;18(6):747-67.

102. Graves N, Weinhold D, Roberts JA. Correcting for bias when estimating the cost of hospital-acquired infection: an analysis of lower respiratory tract infections in non-surgical patients. Health Econ. 2005;14(7):755-61.

103. Hughes RG, Garnick DW, Luft HS, McPhee SJ, Hunt SS. Hospital volume and patient outcomes. The case of hip fracture patients. Med Care. 1988;26(11):1057-67.

104. Farley DE, Ozminkowski RJ. Volume-outcome relationships and in-hospital mortality: the effect of changes in volume over time. Med Care. 1992;30(1):77-94.

105. Picone G, Trogdon J, Jollis J. Hospital volume and quality of care: selective referral or practice makes perfect ? unpublished manuscript. 2005.

106. Gowrisankaran G., Vivian H., Town RJ. Causality and the Volume-Ouctome Relationship in Surgery (unpublished manuscript). In; 2004.

107. Van De Sande S, De Wachter D, Swartenbroekx N, Peers J, Debruyne H, Moldenaers I, et al. Inventaris van databanken gezondheidszorg. Objective Elements - Communication (OEC). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2006. KCE reports 30A (D2006/10.273/14) Available from: http://kce.fgov.be/index_nl.aspx?SGREF=5272&CREF=6845

108. Belgian Cancer Registry. Cancer Incidence in Belgium, 2004-2005. Brussels: 2008. (D/2008/11.846/1) Available from: http://www.coldfusionwebhostings.be/PSK/Upload/GENERAL//Brochures/KIB2004-2005/CancerInc_book.pdf

109. Belgian Cancer Registry. Cancer Incidence in Belgium, 2004-2005 - Appendices. Brussels: 2008. (D/2008/11.846/1) Available from: http://www.coldfusionwebhostings.be/PSK/Upload/GENERAL//Brochures/KIB2004-2005/CancerInc_appendix.pdf

110. Hunninck L. ICD-9-CM Internationale classificatie van ziekten en ingrepen - Codeerhandboek. In: Directoraat-generaal Organisatie gezondheidszorgvoorzieningen Bestuursdirectie Gezondheidszorgbeleid, editor.: FOD Volksgezondheid, Veiligheid van de Voedselketen en Leefmilieu; 2003. p. 181.

111. Gillet P, Kolh P, Sermeus W, Vleugels A, Jacques J, Van den heede K, et al. Detection of adverse events in administrative databases. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE); 2008. D/2008/10.273/75 Available from: http://kce.fgov.be/index_en.aspx?SGREF=5211&CREF=11890

112. Perkins AJ, Kroenke K, Unutzer J, Katon W, Williams JW, Hope C, et al. Common comorbidity scales were similar in their ability to predict health care costs and mortality. J Clin Epidemiol. 2004;57(10):1040-8.

113. Needham DM, Scales DC, Laupacis A, Pronovost PJ. A systematic review of the Charlson comorbidity index using Canadian administrative databases: a perspective on risk adjustment in critical care research. J Crit Care. 2005;20(1):12-9.

114. D'Hoore W, Bouckaert A, Tilquin C. Practical considerations on the use of the Charlson comorbidity index with administrative data bases. J Clin Epidemiol. 1996;49(12):1429-33.

115. Birkmeyer JD, Siewers AE, Finlayson EV, Stukel TA, Lucas FL, Batista I, et al. Hospital volume and surgical mortality in the United States. N Engl J Med. 2002;346(15):1128-37.

116. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med. 2003;349(22):2117-27.

117. Ho V, Heslin MJ. Effect of hospital volume and experience on in-hospital mortality for pancreaticoduodenectomy. Ann Surg. 2003;237(4):509-14.

118. National Cancer Institute. Fact Sheet 5.32: Staging: Questions and Answers. 2004. Available from: http://www.cancer.gov/images/Documents/56fdb7d0-83c4-4ade-8013-11fafc67b3db/fs5_32.pdf

119. College of Oncology. National Clinical Practice Guidelines Oesophageal Cancer. 2008. Available from: https://portal.health.fgov.be/pls/portal/docs/PAGE/INTERNET_PG/HOMEPAGE_MENU/GEZONDHEIDZORG1_MENU/OVERLEGSTRUCTUREN1_MENU/COLLEGESVANGENEESHEREN1_MENU/ONCOLOGIE1_MENU/RICHTLIJNENKLINISCHEPRAKTIJK1_HIDE/RICHTLIJNENKLINISCHEPRAKTIJK1_DOCS/OESOPHAGEAL%20CANCER%20V1.2007%20%28EN%29_1.PDF

120. College of Oncology. National Clinical Practice Guidelines Colon Cancer. 2004. Available from: https://portal.health.fgov.be/pls/portal/docs/PAGE/INTERNET_PG/HOMEPAGE_MENU/GEZONDHEIDZORG1_MENU/OVERLEGSTRUCTUREN1_MENU/COLLEGESVANGENEESHEREN1_MENU/ONCOLOGIE1_MENU/RICHTLIJNENKLINISCHEPRAKTIJK1_HIDE/RICHTLIJNENKLINISCHEPRAKTIJK1_DOCS/COLON%20CANCER_0.PDF

121. Peeters M, Lerut T, Vlayen J, Mambourg F, Ectors N, Deprez P, et al. Wetenschappelijke ondersteuning van het College voor Oncologie: een nationale praktijkrichtlijn voor de aanpak van slokdarm- en maagkanker. Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE). KCE Reports 75A; 2008. (D/2008/10.273/16) Available from: http://kce.fgov.be/index_nl.aspx?SGREF=5263&CREF=10742

122. Dimick JB, Cattaneo SM, Lipsett PA, Pronovost PJ, Heitmiller RF. Hospital volume is related to clinical and economic outcomes of esophageal resection in Maryland. Ann Thorac Surg. 2001;72(2):334-9; discussion 9-41.

123. Gillison EW, Powell J, McConkey CC, Spychal RT. Surgical workload and outcome after resection for carcinoma of the oesophagus and cardia. Br J Surg. 2002;89(3):344-8.

124. Rutegard M, Lagergren J, Rouvelas I, Lagergren P. Surgeon volume is a poor proxy for skill in esophageal cancer surgery. Ann Surg. 2009;249(2):256-61.

125. Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, et al. Chapter 4: Histological groups. In: Cancer Incidence in Five Continents, Vol. IX. IARC Scientific Publications No. 160. Lyon: IARC; 2007. Available from: http://www-dep.iarc.fr/CI5-IX/PDF/chapter4.pdf

126. The Leapfrog Group. Evidence-based Hospital Referral (EBHR) - Factsheet. Washington, DC: 2008. Available from: http://www.leapfroggroup.org/media/file/Leapfrog-Evidence-Based_Hospital_Referral_Fact_Sheet.pdf

127. Rouvelas I, Lindblad M, Zeng W, Viklund P, Ye W, Lagergren J. Impact of hospital volume on long-term survival after esophageal cancer surgery. Arch Surg. 2007;142(2):113-7; discussion 8.

128. Bachmann MO, Alderson D, Edwards D, Wotton S, Bedford C, Peters TJ, et al. Cohort study in South and West England of the influence of specialization on the management and outcome of patients with oesophageal and gastric cancers. Br J Surg. 2002;89(7):914-22.

129. Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. Jama. 1998;280(20):1747-51.

130. Dimick JB, Pronovost PJ, Cowan JA, Lipsett PA. Surgical volume and quality of care for esophageal resection: do high-volume hospitals have fewer complications? Ann Thorac Surg. 2003;75(2):337-41.

131. Dimick JB, Pronovost PJ, Cowan JA, Jr., Lipsett PA, Stanley JC, Upchurch GR, Jr. Variation in postoperative complication rates after high-risk surgery in the United States. Surgery. 2003;134(4):534-40; discussion 40-1.

132. Dimick JB, Cowan JA, Jr., Ailawadi G, Wainess RM, Upchurch GR, Jr. National variation in operative mortality rates for esophageal resection and the need for quality improvement. Arch Surg. 2003;138(12):1305-9.

133. Finlayson EV, Goodney PP, Birkmeyer JD. Hospital volume and operative mortality in cancer surgery: a national study. Arch Surg. 2003;138(7):721-5; discussion 6.

134. Finlayson EV, Birkmeyer JD. Effects of hospital volume on life expectancy after selected cancer operations in older adults: a decision analysis. J Am Coll Surg. 2003;196(3):410-7.

135. Gordon TA, Bowman HM, Bass EB, Lillemoe KD, Yeo CJ, Heitmiller RF, et al. Complex gastrointestinal surgery: impact of provider experience on clinical and economic outcomes. J Am Coll Surg. 1999;189(1):46-56.

136. Kuo EY, Chang Y, Wright CD. Impact of hospital volume on clinical and economic outcomes for esophagectomy. Ann Thorac Surg. 2001;72(4):1118-24.

137. Matthews HR, Powell DJ, McConkey CC. Effect of surgical experience on the results of resection for oesophageal carcinoma. Br J Surg. 1986;73(8):621-3.

138. Milne AA, Skinner J, Browning G. Centralisation of oesophageal cancer services; the view from the periphery. J R Coll Surg Edinb. 2000;45(3):164-7.

139. Patti MG, Corvera CU, Glasgow RE, Way LW. A hospital's annual rate of esophagectomy influences the operative mortality rate. J Gastrointest Surg. 1998;2(2):186-92.

140. Swisher SG, Deford L, Merriman KW, Walsh GL, Smythe R, Vaporicyan A, et al. Effect of operative volume on morbidity, mortality, and hospital use after esophagectomy for cancer. J Thorac Cardiovasc Surg. 2000;119(6):1126-32.

141. Urbach DR, Bell CM, Austin PC. Differences in operative mortality between high- and low-volume hospitals in Ontario for 5 major surgical procedures: estimating the number of lives potentially saved through regionalization. Cmaj. 2003;168(11):1409-14.

142. Urbach DR, Baxter NN. Does it matter what a hospital is "high volume" for? Specificity of hospital volume-outcome associations for surgical procedures: analysis of administrative data. Qual Saf Health Care. 2004;13(5):379-83.

143. van Lanschot JJ, Hulscher JB, Buskens CJ, Tilanus HW, ten Kate FJ, Obertop H. Hospital volume and hospital mortality for esophagectomy. Cancer. 2001;91(8):1574-8.

144. Wenner J, Zilling T, Bladström A, Alvegard T. Influence of surgical volume on hospital mortality for cancer of the esophagus and gastric cardia: a national survey in Sweden 1987-96. In: Proceedings of Book of Abstracts from the European Surgical Association 10th Annual Meeting; 2003; Paris; p. 36-7.

145. Kazui T, Osada H, Fujita H. An attempt to analyze the relation between hospital surgical volume and clinical outcome. Gen Thorac Cardiovasc Surg. 2007;55(12):483-92.

146. Meguid RA, Weiss ES, Chang DC, Brock MV, Yang SC. The effect of volume on esophageal cancer resections: what constitutes acceptable resection volumes for centers of excellence? J Thorac Cardiovasc Surg. 2009;137(1):23-9.

147. Migliore M, Choong CK, Lim E, Goldsmith KA, Ritchie A, Wells FC. A surgeon's case volume of oesophagectomy for cancer strongly influences the operative mortality rate. Eur J Cardiothorac Surg. 2007;32(2):375-80.

148. Rouvelas I, Jia C, Viklund P, Lindblad M, Lagergren J. Surgeon volume and postoperative mortality after oesophagectomy for cancer. Eur J Surg Oncol. 2007;33(2):162-8.

149. Rutegard M, Lagergren P. No influence of surgical volume on patients' health-related quality of life after esophageal cancer resection. Ann Surg Oncol. 2008;15(9):2380-7.

150. Thompson AM, Rapson T, Gilbert FJ, Park KG. Hospital volume does not influence long-term survival of patients undergoing surgery for oesophageal or gastric cancer. Br J Surg. 2007;94(5):578-84.

151. Wenner J, Zilling T, Bladstrom A, Alvegard TA. The influence of surgical volume on hospital mortality and 5-year survival for carcinoma of the oesophagus and gastric cardia. Anticancer Res. 2005;25(1B):419-24.

152. Wouters MW, Wijnhoven BP, Karim-Kos HE, Blaauwgeers HG, Stassen LP, Steup WH, et al. High-volume versus low-volume for esophageal resections for cancer: the essential role of case-mix adjustments based on clinical data. Ann Surg Oncol. 2008;15(1):80-7.

153. Birkmeyer JD, Sun Y, Wong SL, Stukel TA. Hospital volume and late survival after cancer surgery. Ann Surg. 2007;245(5):777-83.

154. Peeters M, Vlayen J, Stordeur S, Mambourg F, Boterberg T, de Hemptinne B, et al. Wetenschappelijke ondersteuning van het College voor Oncologie: een nationale praktijkrichtlijn voor de aanpak van pancreaskanker. Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE). KCE Reports 105A; 2009. (D/2009/10.273/10) Available from: http://kce.fgov.be/index_nl.aspx?SGREF=5263&CREF=12744

155. Glasgow RE, Mulvihill SJ. Hospital volume influences outcome in patients undergoing pancreatic resection for cancer. West J Med. 1996;165(5):294-300.

156. Topal B, Van de Sande S, Fieuws S, Penninckx F. Effect of centralization of pancreaticoduodenectomy on nationwide hospital mortality and length of stay. Br J Surg. 2007;94(11):1377-81.

157. Birkmeyer JD, Finlayson SR, Tosteson AN, Sharp SM, Warshaw AL, Fisher ES. Effect of hospital volume on in-hospital mortality with pancreaticoduodenectomy. Surgery. 1999;125(3):250-6.

158. Birkmeyer JD, Warshaw AL, Finlayson SR, Grove MR, Tosteson AN. Relationship between hospital volume and late survival after pancreaticoduodenectomy. Surgery. 1999;126(2):178-83.

159. Edge SB, Schmieg RE, Jr., Rosenlof LK, Wilhelm MC. Pancreas cancer resection outcome in American University centers in 1989-1990. Cancer. 1993;71(11):3502-8.

160. Goodney PP, Stukel TA, Lucas FL, Finlayson EV, Birkmeyer JD. Hospital volume, length of stay, and readmission rates in high-risk surgery. Ann Surg. 2003;238(2):161-7.

161. Gordon TA, Burleyson GP, Tielsch JM, Cameron JL. The effects of regionalization on cost and outcome for one general high-risk surgical procedure. Ann Surg. 1995;221(1):43-9.

162. Gouma DJ, van Geenen RC, van Gulik TM, de Haan RJ, de Wit LT, Busch OR, et al. Rates of complications and death after pancreaticoduodenectomy: risk factors and the impact of hospital volume. Ann Surg. 2000;232(6):786-95.

163. Imperato PJ, Nenner RP, Starr HA, Will TO, Rosenberg CR, Dearie MB. The effects of regionalization on clinical outcomes for a high risk surgical procedure: a study of the Whipple procedure in New York State. Am J Med Qual. 1996;11(4):193-7.

164. Kotwall CA, Maxwell JG, Brinker CC, Koch GG, Covington DL. National estimates of mortality rates for radical pancreaticoduodenectomy in 25,000 patients. Ann Surg Oncol. 2002;9(9):847-54.

165. Lieberman MD, Kilburn H, Lindsey M, Brennan MF. Relation of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy. Ann Surg. 1995;222(5):638-45.

166. Neoptolemos JP, Russell RC, Bramhall S, Theis B. Low mortality following resection for pancreatic and periampullary tumours in 1026 patients: UK survey of specialist pancreatic units. UK Pancreatic Cancer Group. Br J Surg. 1997;84(10):1370-6.

167. Nordback L, Parviainen M, Raty S, Kuivanen H, Sand J. Resection of the head of the pancreas in Finland: effects of hospital and surgeon on short-term and long-term results. Scand J Gastroenterol. 2002;37(12):1454-60.

168. Simunovic M, To T, Theriault M, Langer B. Relation between hospital surgical volume and outcome for pancreatic resection for neoplasm in a publicly funded health care system. Cmaj. 1999;160(5):643-8.

169. Sosa JA, Bowman HM, Gordon TA, Bass EB, Yeo CJ, Lillemoe KD, et al. Importance of hospital volume in the overall management of pancreatic cancer. Ann Surg. 1998;228(3):429-38.

170. Wade TP, Halaby IA, Stapleton DR, Virgo KS, Johnson FE. Population-based analysis of treatment of pancreatic cancer and Whipple resection: Department of Defense hospitals, 1989-1994. Surgery. 1996;120(4):680-5; discussion 6-7.

171. Yeo CJ, Cameron JL, Maher MM, Sauter PK, Zahurak ML, Talamini MA, et al. A prospective randomized trial of pancreaticogastrostomy versus pancreaticojejunostomy after pancreaticoduodenectomy. Ann Surg. 1995;222(4):580-8; discussion 8-92.

172. Balzano G, Zerbi A, Capretti G, Rocchetti S, Capitanio V, Di Carlo V. Effect of hospital volume on outcome of pancreaticoduodenectomy in Italy. Br J Surg. 2008;95(3):357-62.

173. Bilimoria KY, Bentrem DJ, Ko CY, Tomlinson JS, Stewart AK, Winchester DP, et al. Multimodality therapy for pancreatic cancer in the U.S. : utilization, outcomes, and the effect of hospital volume. Cancer. 2007;110(6):1227-34.

174. Dasgupta R, Kim PC. Relationship between surgical volume and clinical outcome: should pediatric surgeons be doing pancreaticoduodenectomies? J Pediatr Surg. 2005;40(5):793-6.

175. Fong Y, Gonen M, Rubin D, Radzyner M, Brennan MF. Long-term survival is superior after resection for cancer in high-volume centers. Ann Surg. 2005;242(4):540-4; discussion 4-7.

176. Rosemurgy A, Cowgill S, Coe B, Thomas A, Al-Saadi S, Goldin S, et al. Frequency with which surgeons undertake pancreaticoduodenectomy continues to determine length of stay, hospital charges, and in-hospital mortality. J Gastrointest Surg. 2008;12(3):442-9.

177. Peeters M, Zlotta A, Roucoux F, De Greve J, Van Belle S, Haelterman M, et al. Nationale Richtlijnen van het College voor oncologie: A. Algemeen kader oncologish kwaliteitshandboek B. Wetenschappelijke basis voor klinische paden voor diagnose en behandeling colorectale kanker en testiskanker. Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2006. KCE reports 29A (D2006/10.273/12) Available from: http://kce.fgov.be/index_nl.aspx?SGREF=5272&CREF=6892

178. Harmon JW, Tang DG, Gordon TA, Bowman HM, Choti MA, Kaufman HS, et al. Hospital volume can serve as a surrogate for surgeon volume for achieving excellent outcomes in colorectal resection. Ann Surg. 1999;230(3):404-11; discussion 11-3.

179. Rabeneck L, Davila JA, Thompson M, El-Serag HB. Surgical volume and long-term survival following surgery for colorectal cancer in the Veterans Affairs Health-Care System. Am J Gastroenterol. 2004;99(4):668-75.

180. Vlayen J, Verstreken M, Mertens C, Van Eycken E, Penninckx F. Kwaliteit van rectale kankerzorg – Fase 2: ontwikkeling en test van een set van kwaliteitsindicatoren. Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2008. KCE reports 81A (D/2008/10.273/38) Available from: http://kce.fgov.be/index_nl.aspx?SGREF=5272&CREF=11258

181. Lothian and Borders large bowel cancer project: immediate outcome after surgery. The consultant surgeons and pathologists of the Lothian and Borders Health Boards. Br J Surg. 1995;82(7):888-90.

182. Borowski D, Ratcliffe A, Kelly S. The effect of volume and specialization on short-term outcome in surgery for colorectal cancer. Colorectal Dis. 2004;6 (Suppl. 1):54.

183. Burns LR, Wholey DR. The effects of patient, hospital, and physician characteristics on length of stay and mortality. Med Care. 1991;29(3):251-71.

184. Callahan MA, Christos PJ, Gold HT, Mushlin AI, Daly JM. Influence of surgical subspecialty training on in-hospital mortality for gastrectomy and colectomy patients. Ann Surg. 2003;238(4):629-36; discussion 36-9.

185. Dimick JB, Cowan JA, Jr., Upchurch GR, Jr., Colletti LM. Hospital volume and surgical outcomes for elderly patients with colorectal cancer in the United States. J Surg Res. 2003;114(1):50-6.

186. Engel A, Oomen J, Cuesta M. Surgeons with high volume case load make a difference in acute high risk colorectal patients. Colorectal Dis. 2003;5 (Suppl. 1):30.

187. Flood AB, Scott WR, Ewy W. Does practice make perfect? Part I: The relation between hospital volume and outcomes for selected diagnostic categories. Med Care. 1984;22(2):98-114.

188. Hannan EL, O'Donnell JF, Kilburn H, Jr., Bernard HR, Yazici A. Investigation of the relationship between volume and mortality for surgical procedures performed in New York State hospitals. Jama. 1989;262(4):503-10.

189. Hannan EL, Radzyner M, Rubin D, Dougherty J, Brennan MF. The influence of hospital and surgeon volume on in-hospital mortality for colectomy, gastrectomy, and lung lobectomy in patients with cancer. Surgery. 2002;131(1):6-15.

190. Hermanek P, Mansmann U, Staimmer DS, Riedl S, Hermanek P. The German experience: the surgeon as a prognostic factor in colon and rectal cancer surgery. Surg Oncol Clin N Am. 2000;9(1):33-49, vi.

191. Hermanek P, Wiebelt H, Staimmer D, Riedl S. Prognostic factors of rectum carcinoma--experience of the German Multicentre Study SGCRC. German Study Group Colo-Rectal Carcinoma. Tumori. 1995;81(3 Suppl):60-4.

192. Hughes RG, Hunt SS, Luft HS. Effects of surgeon volume and hospital volume on quality of care in hospitals. Med Care. 1987;25(6):489-503.

193. Jessup JM, McGinnis LS, Steele GD, Jr., Menck HR, Winchester DP. The National Cancer Data Base. Report on colon cancer. Cancer. 1996;78(4):918-26.

194. Jibawi A, Ballal M, Willmott S, Selvachandran S, Cade D. Are workload and mortality suitable performance indicators in colorectal cancer? Colorectal Dis. 2004;6 (Suppl. 1):59-60.

195. Kee F, Wilson RH, Harper C, Patterson CC, McCallion K, Houston RF, et al. Influence of hospital and clinician workload on survival from colorectal cancer: cohort study. Bmj. 1999;318(7195):1381-5.

196. Kelly JV, Hellinger FJ. Physician and hospital factors associated with mortality of surgical patients. Med Care. 1986;24(9):785-800.

197. Khuri SF, Daley J, Henderson W, Hur K, Hossain M, Soybel D, et al. Relation of surgical volume to outcome in eight common operations: results from the VA National Surgical Quality Improvement Program. Ann Surg. 1999;230(3):414-29; discussion 29-32.

198. Ko CY, Chang JT, Chaudhry S, Kominski G. Are high-volume surgeons and hospitals the most important predictors of in-hospital outcome for colon cancer resection? Surgery. 2002;132(2):268-73.

199. Maerki SC, Luft HS, Hunt SS. Selecting categories of patients for regionalization. Implications of the relationship between volume and outcome. Med Care. 1986;24(2):148-58.

200. Marusch F, Koch A, Schmidt U, Zippel R, Lehmann M, Czarnetzki HD, et al. Effect of caseload on the short-term outcome of colon surgery: results of a multicenter study. Int J Colorectal Dis. 2001;16(6):362-9.

201. McArdle CS, Hole DJ. Influence of volume and specialization on survival following surgery for colorectal cancer. Br J Surg. 2004;91(5):610-7.

202. McArdle CS, Hole D. Impact of variability among surgeons on postoperative morbidity and mortality and ultimate survival. Bmj. 1991;302(6791):1501-5.

203. Mella J, Biffin A, Radcliffe AG, Stamatakis JD, Steele RJ. Population-based audit of colorectal cancer management in two UK health regions. Colorectal Cancer Working Group, Royal College of Surgeons of England Clinical Epidemiology and Audit Unit. Br J Surg. 1997;84(12):1731-6.

204. Meyerhardt JA, Catalano PJ, Schrag D, Ayanian JZ, Haller DG, Mayer RJ, et al. Association of hospital procedure volume and outcomes in patients with colon cancer at high risk for recurrence. Ann Intern Med. 2003;139(8):649-57.

205. Parry JM, Collins S, Mathers J, Scott NA, Woodman CB. Influence of volume of work on the outcome of treatment for patients with colorectal cancer. Br J Surg. 1999;86(4):475-81.

206. Prystowsky JB, Bordage G, Feinglass JM. Patient outcomes for segmental colon resection according to surgeon's training, certification, and experience. Surgery. 2002;132(4):663-70; discussion 70-2.

207. Riley G, Lubitz J. Outcomes of surgery among the Medicare aged: surgical volume and mortality. Health Care Financ Rev. 1985;7(1):37-47.

208. Sagar PM, Hartley MN, MacFie J, Taylor BA, Copeland GP. Comparison of individual surgeon's performance. Risk-adjusted analysis with POSSUM scoring system. Dis Colon Rectum. 1996;39(6):654-8.

209. Schrag D, Cramer LD, Bach PB, Cohen AM, Warren JL, Begg CB. Influence of hospital procedure volume on outcomes following surgery for colon cancer. Jama. 2000;284(23):3028-35.

210. Schrag D, Panageas KS, Riedel E, Hsieh L, Bach PB, Guillem JG, et al. Surgeon volume compared to hospital volume as a predictor of outcome following primary colon cancer resection. J Surg Oncol. 2003;83(2):68-78; discussion -9.

211. Wrigley H, Roderick P, George S, Smith J, Mullee M, Goddard J. Inequalities in survival from colorectal cancer: a comparison of the impact of deprivation, treatment, and host factors on observed and cause specific survival. J Epidemiol Community Health. 2003;57(4):301-9.

212. Zingmond D, Maggard M, O'Connell J, Liu J, Etzioni D, Ko C. What predicts serious complications in colorectal cancer resection? Am Surg. 2003;69(11):969-74.

213. Beart RW, Steele GD, Jr., Menck HR, Chmiel JS, Ocwieja KE, Winchester DP. Management and survival of patients with adenocarcinoma of the colon and rectum: a national survey of the Commission on Cancer. J Am Coll Surg. 1995;181(3):225-36.

214. Billingsley KG, Morris AM, Dominitz JA, Matthews B, Dobie S, Barlow W, et al. Surgeon and hospital characteristics as predictors of major adverse outcomes following colon cancer surgery: understanding the volume-outcome relationship. Arch Surg. 2007;142(1):23-31; discussion 2.

215. Kuhry E, Bonjer HJ, Haglind E, Hop WC, Veldkamp R, Cuesta MA, et al. Impact of hospital case volume on short-term outcome after laparoscopic operation for colonic cancer. Surg Endosc. 2005;19(5):687-92.

216. Larson DW, Marcello PW, Larach SW, Wexner SD, Park A, Marks J, et al. Surgeon volume does not predict outcomes in the setting of technical credentialing: results from a randomized trial in colon cancer. Ann Surg. 2008;248(5):746-50.

217. Morris M, Platell CF. Surgical volume influences survival in patients undergoing resections for stage II colon cancers. ANZ J Surg. 2007;77(10):902-6.

218. Porter GA, Soskolne CL, Yakimets WW, Newman SC. Surgeon-related factors and outcome in rectal cancer. Ann Surg. 1998;227(2):157-67.

219. Guller U, Safford S, Pietrobon R, Heberer M, Oertli D, Jain NB. High hospital volume is associated with better outcomes for breast cancer surgery: analysis of 233,247 patients. World J Surg. 2005;29(8):994-9; discussion 9-1000.

220. Roohan PJ, Bickell NA, Baptiste MS, Therriault GD, Ferrara EP, Siu AL. Hospital volume differences and five-year survival from breast cancer. Am J Public Health. 1998;88(3):454-7.

221. Stefoski Mikeljevic J, Haward RA, Johnston C, Sainsbury R, Forman D. Surgeon workload and survival from breast cancer. Br J Cancer. 2003;89(3):487-91.

222. Bailie K, Dobie I, Kirk S, Donnelly M. Survival after breast cancer treatment: the impact of provider volume. J Eval Clin Pract. 2007;13(5):749-57.

223. 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):1958-63.

224. Christiaens M-R, Vlayen J, Gailly J, Neven P, Carly B, Schobbens J-C, et al. Wetenschappelijke ondersteuning van het College voor Oncologie: een nationale praktijkrichtlijn voor de aanpak van borstkanker. Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2007. KCE reports 63A (D2007/10.273/35) Available from: http://kce.fgov.be/index_nl.aspx?SGREF=5263&CREF=9623

225. College of Oncology. National Clinical Practice Guidelines Breast Cancer. 2007. Available from: https://portal.health.fgov.be/pls/portal/docs/PAGE/INTERNET_PG/HOMEPAGE_MENU/GEZONDHEIDZORG1_MENU/OVERLEGSTRUCTUREN1_MENU/COLLEGESVANGENEESHEREN1_MENU/ONCOLOGIE1_MENU/RICHTLIJNENKLINISCHEPRAKTIJK1_HIDE/RICHTLIJNENKLINISCHEPRAKTIJK1_DOCS/BREAST%20CANCER%20V1.2007%20(EN)_9.PDF

226. Boffetta P, Merletti F, Winkelmann R, Magnani C, Cappa AP, Terracini B. Survival of breast cancer patients from Piedmont, Italy. Cancer Causes Control. 1993;4(3):209-15.

227. Harcourt KF, Hicks KL. Is there a relationship between case volume and survival in breast cancer? Am J Surg. 2003;185(5):407-10.

228. Polednak AP. Hospital volume and survival of breast cancer patients in Connecticut. Am J Public Health. 1999;89(6):946-7.

229. Sainsbury R, Haward B, Rider L, Johnston C, Round C. Influence of clinician workload and patterns of treatment on survival from breast cancer. Lancet. 1995;345(8960):1265-70.

230. Sloan FA, Perrin JM, Valvona J. In-hospital mortality of surgical patients: is there an empiric basis for standard setting? Surgery. 1986;99(4):446-54.

231. Bouche G, Migeot V, Mathoulin-Pelissier S, Salamon R, Ingrand P. Breast cancer surgery: do all patients want to go to high-volume hospitals? Surgery. 2008;143(6):699-705.

232. Ingram DM, McEvoy SP, Byrne MJ, Fritschi L, Joseph DJ, Jamrozik K. Surgical caseload and outcomes for women with invasive breast cancer treated in Western Australia. Breast. 2005;14(1):11-7.

233. Nomura E, Tsukuma H, Ajiki W, Ishikawa O, Oshima A. Population-based study of the relationship between hospital surgical volume and 10-year survival of breast cancer patients in Osaka, Japan. Cancer Sci. 2006;97(7):618-22.

234. 26 APRIL 2007. - Koninklijk besluit houdende vaststelling van de normen waaraan het gespecialiseerd oncologisch zorgprogramma voor borstkanker moet voldoen om te worden erkend, B.S. 20 juli 2007.

235. Remacle A, Mertens R. Borstkanker. Brussel: februari 2007. Themadossier CM Available from: http://www.cm.be/cm-tridion/nl/100/Resources/CM%20Dossier%20Borstkanker%20def_tcm24-14394.pdf

236. Bach PB, Cramer LD, Schrag D, Downey RJ, Gelfand SE, Begg CB. The influence of hospital volume on survival after resection for lung cancer. N Engl J Med. 2001;345(3):181-8.

237. Sioris T, Sihvo E, Sankila R, Salo J. Effect of surgical volume and hospital type on outcome in non-small cell lung cancer surgery: a Finnish population-based study. Lung Cancer. 2008;59(1):119-25.

238. Lababede O, Meziane MA, Rice TW. TNM staging of lung cancer: a quick reference chart. Chest. 1999;115(1):233-5.

239. Romano PS, Mark DH. Patient and hospital characteristics related to in-hospital mortality after lung cancer resection. Chest. 1992;101(5):1332-7.

240. Silvestri GA, Handy J, Lackland D, Corley E, Reed CE. Specialists achieve better outcomes than generalists for lung cancer surgery. Chest. 1998;114(3):675-80.

241. Freixinet JL, Julia-Serda G, Rodriguez PM, Santana NB, de Castro FR, Fiuza MD, et al. Hospital volume: operative morbidity, mortality and survival in thoracotomy for lung cancer. A Spanish multicenter study of 2994 cases. Eur J Cardiothorac Surg. 2006;29(1):20-5.

242. Li WW, de Mol BA. Hospital volume as a quality standard in lung cancer surgery: an unfinished debate. Ann Thorac Surg. 2008;85(5):1840-1; author reply 1-2.

243. Lien YC, Huang MT, Lin HC. Association between surgeon and hospital volume and in-hospital fatalities after lung cancer resections: the experience of an Asian country. Ann Thorac Surg. 2007;83(5):1837-43.

244. Osada H, Yamakoshi E. Hospital volume and surgical outcomes of lung cancer in Japan. Gen Thorac Cardiovasc Surg. 2007;55(9):360-5.

245. National Institute for Health and Clinical Excellence. Stroke: the diagnosis and initial management of acute stroke and transient ischaemic attack.(Clinical guideline 68). 2008. Available from: http://www.nice.org.uk/nicemedia/pdf/CG68FullGuideline.pdf

246. Jacques J, Gillain D, Fecher F, Van De Sande S, Vrijens F, Ramaekers D, et al. Studie naar praktijkverschillen bij electieve chirurgische ingrepen in België. Good Clinical Practice (GCP). Brussel: Federaal Kenniscentrum voor de gezondheidszorg (KCE); 2006. KCE reports 42A (D/2006/10.273/45)

247. Bonneux L, Cleemput I, Ramaekers D. Endovasculaire behandeling van Carotisstenose. Health Technology Assessment (HTA). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2005 April. KCE reports 13A (D/2005/10.273/09) Available from: http://kce.fgov.be/Download.aspx?ID=264

248. Agency for Healthcare Research and Quality Rockville, MD.;c February 2006. Inpatient Quality Indicators Overview. AHRQ Quality Indicators. Available from: http://www.qualityindicators.ahrq.gov/iqi_overview.htm

249. The Leapfrog Group [cited 21/10/2008]. The Leapfrog Group. Available from: http://www.leapfroggroup.org/
250. Personal Communication by Barbara A. Rudolph, Director Leaps and Measures, The Leapfrog Group. In; 24/10/2008.
251. Roddy SP, O'Donnell TF, Jr., Wilson AL, Estes JM, Mackey WC. The Balanced Budget Act: potential implications for the practice of vascular surgery. J Vasc Surg. 2000;31(2):227-36.
252. Peck C, Peck J, Peck A. Comparison of carotid endarterectomy at high- and low-volume hospitals. Am J Surg. 2001;181(5):450-3.
253. O'Neill L, Lanska DJ, Hartz A. Surgeon characteristics associated with mortality and morbidity following carotid endarterectomy. Neurology. 2000;55(6):773-81.
254. Middleton S, Donnelly N. Outcomes of carotid endarterectomy: how does the Australian state of New South Wales compare with international benchmarks? J Vasc Surg. 2002;36(1):62-9.
255. Harthun NL, Kongable GL, Baglioni AJ, Meakem TD, Kron IL. Examination of sex as an independent risk factor for adverse events after carotid endarterectomy. J Vasc Surg. 2005;41(2):223-30.
256. Hannan EL, Popp AJ, Feustel P, Halm E, Bernardini G, Waldman J, et al. Association of surgical specialty and processes of care with patient outcomes for carotid endarterectomy. Stroke. 2001;32(12):2890-7.
257. Feasby TE, Quan H, Ghali WA. Hospital and surgeon determinants of carotid endarterectomy outcomes. Arch Neurol. 2002;59(12):1877-81.
258. Dardik A, Bowman HM, Gordon TA, Hsieh G, Perler BA. Impact of race on the outcome of carotid endarterectomy: a population-based analysis of 9,842 recent elective procedures. Ann Surg. 2000;232(5):704-9.
259. Cowan JA, Jr., Dimick JB, Thompson BG, Stanley JC, Upchurch GR, Jr. Surgeon volume as an indicator of outcomes after carotid endarterectomy: an effect independent of specialty practice and hospital volume. J Am Coll Surg. 2002;195(6):814-21.
260. Mas JL, Chatellier G, Beyssen B, Branchereau A, Moulin T, Becquemin JP, et al. Endarterectomy versus stenting in patients with symptomatic severe carotid stenosis. N Engl J Med. 2006;355(16):1660-71.
261. Yadav JS, Wholey MH, Kuntz RE, Fayad P, Katzen BT, Mishkel GJ, et al. Protected carotid-artery stenting versus endarterectomy in high-risk patients. N Engl J Med. 2004;351(15):1493-501.
262. American Heart Association [cited 30/10/2008]. Surgery and Other Medical Procedures. Available from: http://www.americanheart.org/presenter.jhtml?identifier=123
263. Sellke FW, DiMaio JM, Caplan LR, Ferguson TB, Gardner TJ, Hiratzka LF, et al. Comparing on-pump and off-pump coronary artery bypass grafting: numerous studies but few conclusions: a scientific statement from the American Heart Association council on cardiovascular surgery and anesthesia in collaboration with the interdisciplinary working group on quality of care and outcomes research. Circulation. 2005;111(21):2858-64.
264. New York State Department of Health [updated 01/10/2008; cited 21/10/2008]. New York State Hospital Profile. Available from: http://hospitals.nyhealth.gov/index.php
265. Neyt M, Van Brabandt H, Devriese S, Mahieu J, De Ridder A, De Graeve D, et al. Drug Eluting Stents in Belgium: Health Technology Assessment. Brussels: Belgian Health Care Knowledge Centre (KCE); 2007. D/2007/10.273/49 Available from: http://kce.fgov.be/index_nl.aspx?SGREF=5272&CREF=10069
266. Shahian DM, Silverstein T, Lovett AF, Wolf RE, Normand S-LT. Comparison of clinical and administrative data sources for hospital coronary artery bypass graft surgery report cards. Circulation. 2007;115(12):1518-27.
267. Brown PP, Mack MJ, Simon AW, Battaglia SL, Tarkington LG, Culler SD, et al. Comparing clinical outcomes in high-volume and low-volume off-pump coronary bypass operation programs. Ann Thorac Surg. 2001;72(3):S1009-15.
268. Nallamothu BK, Saint S, Ramsey SD, Hofer TP, Vijan S, Eagle KA. The role of hospital volume in coronary artery bypass grafting: is more always better? J Am Coll Cardiol. 2001;38(7):1923-30.
269. Carey JS, Robertson JM, Misbach GA, Fisher AL. Relationship of hospital volume to outcome in cardiac surgery programs in California. Am Surg. 2003;69(1):63-8.

270. Glance LG, Dick AW, Mukamel DB, Osler TM. Is the hospital volume-mortality relationship in coronary artery bypass surgery the same for low-risk versus high-risk patients? Ann Thorac Surg. 2003;76(4):1155-62.

271. Hannan EL, Wu C, Ryan TJ, Bennett E, Culliford AT, Gold JP, et al. Do hospitals and surgeons with higher coronary artery bypass graft surgery volumes still have lower risk-adjusted mortality rates? Circulation. 2003;108(7):795-801.

272. Peterson ED, Coombs LP, DeLong ER, Haan CK, Ferguson TB. Procedural volume as a marker of quality for CABG surgery. Jama. 2004;291(2):195-201.

273. Wu C, Hannan EL, Ryan TJ, Bennett E, Culliford AT, Gold JP, et al. Is the impact of hospital and surgeon volumes on the in-hospital mortality rate for coronary artery bypass graft surgery limited to patients at high risk? Circulation. 2004;110(7):784-9.

274. Gammie JS, O'Brien SM, Griffith BP, Ferguson TB, Peterson ED. Influence of hospital procedural volume on care process and mortality for patients undergoing elective surgery for mitral regurgitation. Circulation. 2007;115(7):881-7.

275. Ricciardi R, Virnig BA, Ogilvie Jr JW, Dahlberg PS, Selker HP, Baxter NN. Volume-outcome relationship for coronary artery bypass grafting in an era of decreasing volume. Archives of Surgery. 2008;143(4):338-44.

276. Ad N, Barnett SD, Speir AM. The performance of the EuroSCORE and the Society of Thoracic Surgeons mortality risk score: the gender factor. Interact Cardiovasc Thorac Surg. 2007;6(2):192-5.

277. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg. 1999;16(1):9-13.

278. American Heart Association [cited 29/10/2008]. Percutaneous Coronary Interventions. Available from: http://www.americanheart.org/presenter.jhtml?identifier=4454

279. Smith SC, Jr., Dove JT, Jacobs AK, Kennedy JW, Kereiakes D, Kern MJ, et al. ACC/AHA guidelines of percutaneous coronary interventions (revision of the 1993 PTCA guidelines)--executive summary. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (committee to revise the 1993 guidelines for percutaneous transluminal coronary angioplasty). J Am Coll Cardiol. 2001;37(8):2215-39.

280. American Heart Association [cited 29/10/2008]. Stent Procedure. Available from: http://www.americanheart.org/presenter.jhtml?identifier=4721

281. Jollis JG, Peterson ED, Nelson CL, Stafford JA, DeLong ER, Muhlbaier LH, et al. Relationship between physician and hospital coronary angioplasty volume and outcome in elderly patients. Circulation. 1997;95(11):2485-91.

282. Carey JS, Danielsen B, Gold JP, Rossiter SJ. Procedure rates and outcomes of coronary revascularization procedures in California and New York. J Thorac Cardiovasc Surg. 2005;129(6):1276-82.

283. Epstein AJ, Rathore SS, Krumholz HM, Volpp KG. Volume-based referral for cardiovascular procedures in the United States: a cross-sectional regression analysis. BMC Health Serv Res. 2005;5(1):42.

284. Hannan EL, Wu C, Walford G, King SB, 3rd, Holmes DR, Jr., Ambrose JA, et al. Volume-outcome relationships for percutaneous coronary interventions in the stent era. Circulation. 2005;112(8):1171-9.

285. Moscucci M, Share D, Smith D, O'Donnell MJ, Riba A, McNamara R, et al. Relationship between operator volume and adverse outcome in contemporary percutaneous coronary intervention practice: an analysis of a quality-controlled multicenter percutaneous coronary intervention clinical database. J Am Coll Cardiol. 2005;46(4):625-32.

286. Mukherjee D, Wainess RM, Dimick JB, Cowan JA, Rajagopalan S, Chetcuti S, et al. Variation in outcomes after percutaneous coronary intervention in the United States and predictors of periprocedural mortality. Cardiology. 2005;103(3):143-7.

287. Zahn R, Vogt A, Zeymer U, Gitt AK, Seidl K, Gottwik M, et al. In-hospital time to treatment of patients with acute ST elevation myocardial infarction treated with primary angioplasty: determinants and outcome. Results from the registry of percutaneous coronary interventions in acute myocardial infarction of the Arbeitsgemeinschaft Leitender Kardiologischer Krankenhausarzte. Heart. 2005;91(8):1041-6.

288. Epstein AJ, Rathore SS, Volpp KG, Krumholz HM. Hospital percutaneous coronary intervention volume and patient mortality, 1998 to 2000: does the evidence support current procedure volume minimums? J Am Coll Cardiol. 2004;43(10):1755-62.

289. Harjai KJ, Berman AD, Grines CL, Kahn J, Marsalese D, Mehta RH, et al. Impact of interventionalist volume, experience, and board certification on coronary angioplasty outcomes in the era of stenting. Am J Cardiol. 2004;94(4):421-6.

290. Kansagra SM, Curtis LH, Schulman KA. Regionalization of percutaneous transluminal coronary angioplasty and implications for patient travel distance. Jama. 2004;292(14):1717-23.

291. Rubartelli P, Niccoli L, Alberti A, Giachero C, Ettori F, Missiroli B, et al. Coronary rotational atherectomy in current practice: acute and mid-term results in high- and low-volume centers. Catheter Cardiovasc Interv. 2004;61(4):463-71.

292. Tsuchihashi M, Tsutsui H, Tada H, Shihara M, Takeshita A, Kono S. Volume-outcome relation for hospitals performing angioplasty for acute myocardial infarction: results from the Nationwide Japanese Registry. Circ J. 2004;68(10):887-91.

293. Ward MM, Jaana M, Wakefield DS, Ohsfeldt RL, Schneider JE, Miller T, et al. What would be the effect of referral to high-volume hospitals in a largely rural state? J Rural Health. 2004;20(4):344-54.

294. Nobilio L, Ugolini C. Selective referrals in a 'hub and spoke' institutional setting: the case of coronary angioplasty procedures. Health Policy. 2003;63(1):95-107.

295. Vakili BA, Brown DL. Relation of total annual coronary angioplasty volume of physicians and hospitals on outcomes of primary angioplasty for acute myocardial infarction (data from the 1995 Coronary Angioplasty Reporting System of the New York State Department of Health). Am J Cardiol. 2003;91(6):726-8.

296. Kimmel SE, Sauer WH, Brensinger C, Hirshfeld J, Haber HL, Localio AR. Relationship between coronary angioplasty laboratory volume and outcomes after hospital discharge. Am Heart J. 2002;143(5):833-40.

297. Vakili BA, Kaplan R, Brown DL. Volume-outcome relation for physicians and hospitals performing angioplasty for acute myocardial infarction in New York state. Circulation. 2001;104(18):2171-6.

298. Cannon CP, Gibson CM, Lambrew CT, Shoultz DA, Levy D, French WJ, et al. Relationship of symptom-onset-to-balloon time and door-to-balloon time with mortality in patients undergoing angioplasty for acute myocardial infarction. Jama. 2000;283(22):2941-7.

299. Canto JG, Every NR, Magid DJ, Rogers WJ, Malmgren JA, Frederick PD, et al. The volume of primary angioplasty procedures and survival after acute myocardial infarction. National Registry of Myocardial Infarction 2 Investigators. N Engl J Med. 2000;342(21):1573-80.

300. Every NR, Maynard C, Schulman K, Ritchie JL. The association between institutional primary angioplasty procedure volume and outcome in elderly Americans. J Invasive Cardiol. 2000;12(6):303-8.

301. Ho V. Evolution of the volume-outcome relation for hospitals performing coronary angioplasty. Circulation. 2000;101(15):1806-11.

302. Magid DJ, Calonge BN, Rumsfeld JS, Canto JG, Frederick PD, Every NR, et al. Relation between hospital primary angioplasty volume and mortality for patients with acute MI treated with primary angioplasty vs thrombolytic therapy. Jama. 2000;284(24):3131-8.

303. Maynard C, Every NR, Chapko MK, Ritchie JL. Outcomes of coronary angioplasty procedures performed in rural hospitals. Am J Med. 2000;108(9):710-3.

304. McGrath PD, Wennberg DE, Dickens JD, Jr., Siewers AE, Lucas FL, Malenka DJ, et al. Relation between operator and hospital volume and outcomes following percutaneous coronary interventions in the era of the coronary stent. Jama. 2000;284(24):3139-44.

305. Killip T, 3rd, Kimball JT. Treatment of myocardial infarction in a coronary care unit. A two year experience with 250 patients. Am J Cardiol. 1967;20(4):457-64.

306. Heart Failure Society of America [updated 29/09/2006 cited 20/10/2008]. The Stages of Heart Failure – NYHA Classification. Available from: http://www.abouthf.org/questions_stages.htm

307. Krakauer H, Shekar SS, Kaye MP. The relationship of clinical outcomes to status as a Medicare-approved heart transplant center. Transplantation. 1995;59(6):840-6.

308. Hosenpud JD, Breen TJ, Edwards EB, Daily OP, Hunsicker LG. The effect of transplant center volume on cardiac transplant outcome. A report of the United Network for Organ Sharing Scientific Registry. Jama. 1994;271(23):1844-9.

309. Taylor HD, Dennis DA, Crane HS. Relationship between mortality rates and hospital patient volume for Medicare patients undergoing major orthopaedic surgery of the hip, knee, spine, and femur. J Arthroplasty. 1997;12(3):235-42.

310. Vlayen J, Van De Water G, Camberlin C, Paulus D, Leys M, Ramaekers D, et al. Clinical Quality Indicators. Objective Elements - Communication (OEC). Brussels: Belgian Health Care Knowledge Centre (KCE); 2006. D/2006/10.273/43 Available from: http://kce.fgov.be/index_en.aspx?SGREF=5223&CREF=9140

311. Katz JN, Losina E, Barrett J, Phillips CB, Mahomed NN, Lew RA, et al. Association between hospital and surgeon procedure volume and outcomes of total hip replacement in the United States medicare population. J Bone Joint Surg Am. 2001;83-A(11):1622-9.

312. Diels J, Mertens R, Boly J, de Béthune X, Hutsebaut L, Van den Oever R. Prothèse totale de la hanche - Variations des pratiques médicales et résultats à long terme. 2000. Dossier thématique des MC N° 2. Available from: http://www.cm.be/cm-tridion/fr/130/Resources/prothese_de_hanche_tcm180-8326.pdf

313. Diels J, Mertens R, Boly J, de Béthune X, Hutsebaut L, Van den Oever R. Totale heupprothese - Variatie in medische praktijk en lange termijn-resultaten. 2000. CM Themadossier Nr. 2. Available from: http://www.mc.be/cm-tridion/nl/108/Resources/Dossier_Heupprothese_tcm252-7533.pdf

314. Losina E, Barrett J, Mahomed NN, Baron JA, Katz JN. Early failures of total hip replacement: effect of surgeon volume. Arthritis Rheum. 2004;50(4):1338-43.

315. Solomon DH, Losina E, Baron JA, Fossel AH, Guadagnoli E, Lingard EA, et al. Contribution of hospital characteristics to the volume-outcome relationship: dislocation and infection following total hip replacement surgery. Arthritis Rheum. 2002;46(9):2436-44.

316. Thompson R, Kane RL, Gromala T, McLaughlin B, Flood S, Morris N, et al. Complications and short-term outcomes associated with total hip arthroplasty in teaching and community hospitals. J Arthroplasty. 2002;17(1):32-40.

317. Culler SD, Holmes AM, Gutierrez B. Expected hospital costs of knee replacement for rural residents by location of service. Med Care. 1995;33(12):1188-209.

318. Espehaug B, Havelin LI, Engesaeter LB, Vollset SE. The effect of hospital-type and operating volume on the survival of hip replacements. A review of 39,505 primary total hip replacements reported to the Norwegian Arthroplasty Register, 1988-1996. Acta Orthop Scand. 1999;70(1):12-8.

319. Kreder HJ, Deyo RA, Koepsell T, Swiontkowski MF, Kreuter W. Relationship between the volume of total hip replacements performed by providers and the rates of postoperative complications in the state of Washington. J Bone Joint Surg Am. 1997;79(4):485-94.

320. Kreder HJ, Williams JI, Jaglal S, Hu R, Axcell T, Stephen D. Are complication rates for elective primary total hip arthroplasty in Ontario related to surgeon and hospital volumes? A preliminary investigation. Can J Surg. 1998;41(6):431-7.

321. Lavernia CJ, Guzman JF. Relationship of surgical volume to short-term mortality, morbidity, and hospital charges in arthroplasty. J Arthroplasty. 1995;10(2):133-40.

322. Katz JN, Phillips CB, Baron JA, Fossel AH, Mahomed NN, Barrett J, et al. Association of hospital and surgeon volume of total hip replacement with functional status and satisfaction three years following surgery. Arthritis Rheum. 2003;48(2):560-8.

323. Zhan C, Kaczmarek R, Loyo-Berrios N, Sangl J, Bright RA. Incidence and short-term outcomes of primary and revision hip replacement in the United States. J Bone Joint Surg Am. 2007;89(3):526-33.

324. Mohamed K, Copeland GP, Boot DA, Casserley HC, Shackleford IM, Sherry PG, et al. An assessment of the POSSUM system in orthopaedic surgery. J Bone Joint Surg Br. 2002;84(5):735-9.

325. Katz JN, Barrett J, Mahomed NN, Baron JA, Wright RJ, Losina E. Association between hospital and surgeon procedure volume and the outcomes of total knee replacement. J Bone Joint Surg Am. 2004;86-A(9):1909-16.

326. Feinglass J, Amir H, Taylor P, Lurie I, Manheim LM, Chang RW. How safe is primary knee replacement surgery? Perioperative complication rates in Northern Illinois, 1993-1999. Arthritis Rheum. 2004;51(1):110-6.

327. Graul TL. Total joint replacement: baseline benchmark data for interdisciplinary outcomes management. Orthop Nurs. 2002;21(3):57-64; quiz 5-7.

328. Hervey SL, Purves HR, Guller U, Toth AP, Vail TP, Pietrobon R. Provider Volume of Total Knee Arthroplasties and Patient Outcomes in the HCUP-Nationwide Inpatient Sample. J Bone Joint Surg Am. 2003;85-A(9):1775-83.

329. Kreder HJ, Grosso P, Williams JI, Jaglal S, Axcell T, Wal EK, et al. Provider volume and other predictors of outcome after total knee arthroplasty: a population study in Ontario. Can J Surg. 2003;46(1):15-22.

330. Benjamin G. Three essays on volume, complications and hospital resource use: the case of knee replacement surgery (PhD): Indiana University; 1995.

331. Coyte PC, Hawker G, Croxford R, Wright JG. Rates of revision knee replacement in Ontario, Canada. J Bone Joint Surg Am. 1999;81(6):773-82.

332. Gutierrez B. Three essays on volume, complications and hospital resource use: the case of knee replacement surgery (PhD): Indiana University; 1995.

333. Gutierrez B, Culler SD, Freund DA. Does hospital procedure-specific volume affect treatment costs? A national study of knee replacement surgery. Health Serv Res. 1998;33(3 Pt 1):489-511.

334. Heck DA, Melfi CA, Mamlin LA, Katz BP, Arthur DS, Dittus RS, et al. Revision rates after knee replacement in the United States. Med Care. 1998;36(5):661-9.

335. Heck DA, Robinson RL, Partridge CM, Lubitz RM, Freund DA. Patient outcomes after knee replacement. Clin Orthop Relat Res. 1998(356):93-110.

336. Norton EC, Garfinkel SA, McQuay LJ, Heck DA, Wright JG, Dittus R, et al. The effect of hospital volume on the in-hospital complication rate in knee replacement patients. Health Serv Res. 1998;33(5 Pt 1):1191-210.

337. AO Foundation [cited 03/12/2008]. Evidence summary on proximal femur fractures. Available from: http://www.aofoundation.org/wps/portal/!ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3hng7BARydDRwML1yBXAyMvYz8zEwNPQwN3A6B8JJK8gUWAm4GRk6m_oUlwgBFlHr9uP4_83FT9gtyIcgCExWfz/dl2/d1/L2dJQSEvUUt3QS9ZQnB3LzZfQzBWUUFCMUEwR0dSNTAySkowOFVIRzIwVDQ!/?segment=Proximal&bone=Femur&soloState=true&popupStyle=diagnosis&contentUrl=srg/popup/evidence_sum/1-proximal-femur-fractures.jsp

338. Hamilton BH, Ho V. Does practice make perfect? Examining the relationship between hospital surgical volume and outcomes for hip fracture patients in Quebec. Med Care. 1998;36(6):892-903.

339. Osteoporose, tweede herziene richtlijn. Kwaliteitsinsituut voor de Gezondheidszorg CBO; 2002. Available from: http://www.cbo.nl/product/richtlijnen/folder20021023121843/rl_overzicht/view#14

340. Wenning M, Hupe K, Scheuer I, Senninger N, Smektala R, Windhorst T. [Does quantity mean quality? An analysis of 116,000 patients regarding the connection between the number of cases and the quality of results]. Chirurg. 2000;71(6):717-22.

341. Shah SN, Wainess RM, Karunakar MA. Hemiarthroplasty for femoral neck fracture in the elderly surgeon and hospital volume-related outcomes. J Arthroplasty. 2005;20(4):503-8.

342. Holt G, Smith R, Duncan K, Hutchison JD, Gregori A. Gender differences in epidemiology and outcome after hip fracture: evidence from the Scottish Hip Fracture Audit. J Bone Joint Surg Br. 2008;90(4):480-3.

343. National Institute for Health and Disability Insurance Brussels;c 2009 [updated 6/05/2009; cited 19/05/2009]. Persbericht: Een vooruitgang in de elektronische registratie van gegevens van heup- en knieprothesen. Available from: http://www.riziv.fgov.be/news/nl/press/pdf/press20090506.pdf

344. Department of Health. Guidance on commissioning cancer services: improving outcomes in upper gastro-intestinal cancers - the manual. 2001. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4010025

345. Siriwardena AK. Centralisation of upper gastrointestinal cancer surgery. Ann R Coll Surg Engl. 2007;89(4):335-6.

346. Maria Hewitt for the Committee on Quality of Health Care in America and the National Cancer Policy Board. Interpreting the Volume-Outcome Relationship in the Context of Health Care Quality: Workshop Summary. Washington, D.C.: Institute of Medicine; 2000.

This page is left intentionally blank.

Legal depot : D/2009/10.273/35

KCE reports

33 Effects and costs of pneumococcal conjugate vaccination of Belgian children. D/2006/10.273/54.
34 Trastuzumab in Early Stage Breast Cancer. D/2006/10.273/25.
36 Pharmacological and surgical treatment of obesity. Residential care for severely obese children in Belgium. D/2006/10.273/30.
37 Magnetic Resonance Imaging. D/2006/10.273/34.
38 Cervical Cancer Screening and Human Papillomavirus (HPV) Testing D/2006/10.273/37.
40 Functional status of the patient: a potential tool for the reimbursement of physiotherapy in Belgium? D/2006/10.273/53.
47 Medication use in rest and nursing homes in Belgium. D/2006/10.273/70.
48 Chronic low back pain. D/2006/10.273.71.
49 Antiviral agents in seasonal and pandemic influenza. Literature study and development of practice guidelines. D/2006/10.273/67.
54 Cost-effectiveness analysis of rotavirus vaccination of Belgian infants D/2007/10.273/11.
59 Laboratory tests in general practice D/2007/10.273/26.
60 Pulmonary Function Tests in Adults D/2007/10.273/29.
64 HPV Vaccination for the Prevention of Cervical Cancer in Belgium: Health Technology Assessment. D/2007/10.273/43.
65 Organisation and financing of genetic testing in Belgium. D/2007/10.273/46.
66. Health Technology Assessment: Drug-Eluting Stents in Belgium. D/2007/10.273/49.
70. Comparative study of hospital accreditation programs in Europe. D/2008/10.273/03
71. Guidance for the use of ophthalmic tests in clinical practice. D/200810.273/06.
72. Physician workforce supply in Belgium. Current situation and challenges. D/2008/10.273/09.
74 Hyperbaric Oxygen Therapy: a Rapid Assessment. D/2008/10.273/15.
76. Quality improvement in general practice in Belgium: status quo or quo vadis? D/2008/10.273/20
82. 64-Slice computed tomography imaging of coronary arteries in patients suspected for coronary artery disease. D/2008/10.273/42
83. International comparison of reimbursement principles and legal aspects of plastic surgery. D/200810.273/45
87. Consumption of physiotherapy and physical and rehabilitation medicine in Belgium. D/2008/10.273/56
90. Making general practice attractive: encouraging GP attraction and retention D/2008/10.273/66.
91 Hearing aids in Belgium: health technology assessment. D/2008/10.273/69.
92. Nosocomial Infections in Belgium, part I: national prevalence study. D/2008/10.273/72.
93. Detection of adverse events in administrative databases. D/2008/10.273/75.
95. Percutaneous heart valve implantation in congenital and degenerative valve disease. A rapid Health Technology Assessment. D/2008/10.273/81
100. Threshold values for cost-effectiveness in health care. D/2008/10.273/96
102. Nosocomial Infections in Belgium: Part II, Impact on Mortality and Costs. D/2009/10.273/03
103 Mental health care reforms: evaluation research of 'therapeutic projects' - first intermediate report. D/2009/10.273/06.
104. Robot-assisted surgery: health technology assessment. D/2009/10.273/09
108. Tiotropium in the Treatment of Chronic Obstructive Pulmonary Disease: Health Technology Assessment. D/2009/10.273/20
109. The value of EEG and evoked potentials in clinical practice. D/2009/10.273/23
111. Pharmaceutical and non-pharmaceutical interventions for Alzheimer's Disease, a rapid assessment. D/2009/10.273/29
112 Policies for Orphan Diseases and Orphan Drugs. D/2009/10.273/32.
113 The volume of surgical interventions and its impact on the outcome: feasibility study based on Belgian data

This list only includes those KCE reports for which a full English version is available. However, all KCE reports are available with a French or Dutch executive summary and often contain a scientific summary in English.