Aus dem Institut für Medizinische Biometrie

Universitätsklinikum Heidelberg

Geschäftsführender Direktor: Prof. Dr. sc. hum. Meinhard Kieser

Impact of Institutional Research Activity on Quality of Care and Patient Outcomes in Ovarian Cancer

Inauguraldissertation
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)
an der
Medizinischen Fakultät Heidelberg
der
Ruprecht-Karls-Universität

vorgelegt von
Justine Rochon

aus
Preußisch Stargard (Polen)

2023

Dekan: Prof. Dr. med. Dr. h.c. Hans-Georg Kräusslich

Doktorvater: Prof. Dr. sc. hum. Meinhard Kieser

Contents

1 Introduction

Clinical research refers to investigations that are conducted in human volunteers. It covers a broad spectrum of activities, ranging from basic biomedical research on the mechanisms of diseases, evaluations of biomarkers, assessments of preventive and diagnostic tools through to observational and experimental studies examining safety and efficacy of treatments and medical interventions.

From a scientific perspective, interventional studies (i.e., clinical trials) are considered essential to the progress of medicine and to improve health care outcomes. They are designed to provide answers to questions about interventions that have not yet clearly demonstrated benefits to patients and to find out whether new treatment approaches are safe and effective. In the long term, ineffective or even harmful treatments can be identified and can be replaced by effective and safe ones. Specifically, properly done randomized controlled trials (RCTs) are widely regarded as the gold standard in clinical research because they provide the highest level of evidence for establishing the efficacy of medical interventions. Moreover, RCTs play an important role in making new treatments available. If a treatment proves effective in an RCT, it may become a new therapeutic standard that can help many patients. Results from such trials often serve as basis for development of treatment guidelines and implementation of new clinical pathways. Indeed, new medical knowledge is generated through clinical trials and most of the treatments used today are the result of past research contributions.

One of the most critical success factors of clinical research is access to participants. On the one hand, there is an obvious need for well-designed and well-conducted studies addressing relevant and so far, unresolved questions. On the other hand, public perceptions of study participation are diverse, and patients as well as health care professionals may have some hopes and expectations but also concerns about participation in clinical research. As with any choice, there are pros and cons. Clinical trials may have possible benefits as well as risks for individuals. In addition, clinical trials are, by definition, experimental. It means that it is not known in advance whether the new treatment has any effect, beneficial or harmful. The premise is, therefore, that trials are not designed to benefit current patients and that clinical research only provides the evidence necessary to improve care of future patients. Of course, trial participants may benefit too because they may get access to breakthrough therapies not yet available elsewhere. Trials may offer them the gains of tomorrow's medicine today. However, the justification for randomly assigning participants is based on clinical equipoise or the “uncertainty principle” (Ashcroft 1999). Equipoise means that a subject can be enrolled into a randomized trial only if there is uncertainty about which of the treatments tested in the trial would be most effective. Consequently, there is no guarantee that a new treatment will work, and even if it is effective, it may not work for a particular patient. Even treatments demonstrated as effective for many patients and declared standard are not helpful to everyone. Furthermore, subjects enrolled in an RCT may not receive the experimental treatment because they are allocated to the control arm. Moreover, a new treatment may have unknown risks that are worse than those of standard treatment are. The conventional view is, therefore, that trial participants sacrifice their time, undertake risks for the good of future generations of patients, and usually do not benefit directly from trial participation. Those volunteers who enroll in trials have thus to understand that they primarily help others and not themselves. When people are asked to give reasons why they participate in clinical research, some of them indeed refer to altruistic motives for participating in trials that offer no direct therapeutic benefit but expose them to significant risks of harm (Godskesen et al. 2015).

Through engagement in clinical research, individuals can contribute to scientific and societal progress. However, although trial participants commonly report that altruism motivated their decision to enroll, it is rarely their primary factor in the decision-making process about participation, especially for patients with serious or life-threatening disease. For example, Truong et al. (2011) identified low levels of altruism in patients with poor prognosis that were enrolled in early-phase cancer trials. For those patients, participation is often driven by hope for a cure (Godskesen et al. 2015). Many patients who join clinical trials are eager to benefit from newest research and are thus optimistic about their trial participation. They often perceive trials as the only way of being treated with innovative therapies and consequently agree to participate in trials to get a chance of clinical improvement that is unlikely with conventional therapy. Many health care professionals have similar optimistic expectations and support clinical research because they believe that trials provide the best available treatment for their patients. Sometimes, clinical trials are even seen as a last resort for patients for whom no other treatment option exists. Especially in the field of oncology, there is a professional view that clinical trials offer superior, most up-to-date care from experts. For example, the National Cancer Institute's (2016) booklet states that "Clinical trials are the key to making progress against cancer" (p. 2); trials facilitate access to "high-quality cancer care" and "If a new treatment is proven to work and you are receiving it, you may be among the first to benefit" (p. 11). In addition, many announcements related to cancer clinical research promise that even if patients enter a study but do not receive the experimental therapy, they will at least receive the available standard therapy which may be as good as, or even better than, the new approach. For this reason, it is not unusual nowadays to see patients with common cancers getting their first treatment in a clinical trial setting. In addition, trials can involve diagnostic tests or procedures that are performed solely for research purposes. Consequently, participation in clinical research may result in a quality of care and close monitoring patients may not otherwise have access to. This makes trials distinct from routine clinical practice. It has thus been suggested that patients in trials, even those treated in control arms, get better treatment and may therefore have better health care outcomes than non-trial patients.

More than a decade ago, a workshop on "Clinical Research and Healthcare Outcomes" was held at the International Agency for Research on Cancer (IARC) in Lyon, France (Selby 2011). This workshop explored the evidence for possible advantages from establishing research-intensive health care systems. A key topic throughout the meeting was the so-called "second translational gap". Usually, there are two main types of translational gaps: The first gap refers to the barriers of moving from basic biomedical research into clinical knowledge, while the second gap refers to the implementation of clinical research results into daily practice in order to improve patient health (Sung et al. 2003; Woolf 2008; Grimshaw et al. 2012). Indeed, one of the most consistent results from literature is the gap between evidence and practice. Because of this gap, many patients fail to receive optimal care or are even harmed because of inappropriate care (Grol and Grimshaw 2003; Grimshaw et al. 2012). For example, Bristow et al. (2013) reported that only 37% of patients diagnosed with ovarian cancer in California between 1999 and 2006 received health care in line with the National Comprehensive Cancer Network guidelines. Similar results have been reported earlier for other indications and care settings (Grol 2001). Consequently, increased efforts to close the second translational gap and to move research results into practice have been made over the last decades. These efforts are known under different terms such as quality assurance, implementation research, knowledge utilization, and knowledge translation, to name just a few of them (McKibbon et al. 2010). Knowledge translation, for example, has been defined by Grimshaw et al. (2012) as "ensuring that stakeholders are aware of and use

research evidence to inform their health and health care decision-making" (p. 2). This definition acknowledges that there are different stakeholders and recipients of knowledge, including health care professionals, researchers, industry sponsors, regulators, health technology assessment bodies, and policy makers, but also the patients and their relatives as the consumers of care. Despite the existing multiple terms, the traditional approach to close the second translational gap has been based up to now on knowledge translation and education. The primary goal has been mainly to disseminate new positive results from individual trials and to inform the health community about their relevance in real world settings. However, it has been recognized that the simple availability of new research results does not automatically lead to their use and that patients often fail to benefit optimally and timely from advances in research. Reading articles and books or attending conferences have generally not proven to be effective in changing professional behavior or routines (Grimshaw and Thomson 1998). According to Morris et al. (2011), it needs an average of 17 years for research results to reach clinical practice. And in many cases, research even does not lead to worthwhile achievements (Chalmers et al. 2014). In consequence, new translational activities have been conceptualized. Among others, it has been proposed to use the conduct of research itself as a bridge for the evidence-practice gap. At the IARC workshop special interest was given to advantages of being involved in clinical research on institutional level. The starting point was the suggestion that clinical research can translate into better health directly by improving the outcomes for individual trial patients and indirectly by improving the health care services in research-active institutions (Selby 2011).

Research effects must be visible and translate into better health care to justify the huge investments it requires. Demonstrating that research is beneficial and that resources are efficiently allocated and used is therefore crucial to get the buy-in from all important stakeholders. Even though the public opinion seems to recognize intuitively that clinical research is worthwhile, the positive link between clinical trials and better outcome often holds as self-evident and many health care providers make claims supporting research activity without corresponding data. In the current culture of opinion-based policy making, this type of thinking seems to have been accepted without really challenging the underlying concept (Krzyzanowska et al. 2011). At the same time, the characteristics of research-active or research-intensive health care systems are poorly understood, and there is an urgent need to increase value from research and to reduce waste in medicine (Moher et al. 2016). Therefore, it is worth to learn more about "trial effects" and to clarify the relevant questions to be addressed.

1.1 Trial effects

For many years the effects of participation in clinical research have been extensively studied and controversially discussed in the literature. In particular, trial effects have been explored in oncology more than in other areas. One reason for that might be that cancer survival serves as a crucial indicator of the effectiveness of health care systems (De Angelis et al. 2014). Cancer survival rates usually reflect the ability of these systems to diagnose, treat, and manage cancer in a timely and efficient manner. High survival rates suggest that the health care system is providing patients with access to advanced diagnostic tools, cutting-edge treatments, and comprehensive care, ultimately leading to better outcomes. Furthermore, monitoring cancer survival rates allows for the identification of disparities and areas for improvement, enabling health care systems to continuously evolve and enhance the quality of care provided to patients. Participation in clinical research is one of the factors being investigated for its effect on standards of care and patient outcomes.

Previous research has mainly focused on the impact on patients and their experiences of being actively involved in clinical trials—the so-called "participation effect" (Selby 2011). Consequently, in most studies, "trial patients" were compared with "non-trial patients" and the question asked was "Do patients in clinical trials do better than patients outside trials (in the same institution or health care service)?" (Selby and Autier 2011, p. 6). In this thesis, I refer to this as Question 1.

A first review evaluating the effect of participation in clinical trials (Stiller 1994) provided weak evidence for the existence of a beneficial trial effect related to Question 1. For example, a study performed by Ward et al. (1992) showed higher survival rates for patients entered in the first British Stomach Cancer Group trial (Fielding et al. 1983) compared to concurrent non-trial patients from the same districts. However, when the comparison was confined to those patients that actually met the eligibility criteria for the trial, such as fitness, the effect mostly disappeared. Similarly, Bertelsen (1991) found significantly higher survival rates for advanced disease cancer patients randomized to trials compared to non-randomized patients. However, the authors noted that chemotherapy was given to all patients who were randomized but only to the half of those who were not—many of the non-randomized patients were simply too ill to start chemotherapy treatment. When the comparison was restricted to patients that actually received chemotherapy, the difference was not significant anymore. In another study included in Stiller's review, Karjalainen and Palva (1989) reported higher survival rates for multiple myeloma patients in districts with local policy to enroll these patients in clinical trials compared to their peers from non-trial districts. The authors wondered about what might have caused this effect and suggested that patients in research-active areas benefited from the use of trial protocols. Braunholtz and colleagues (2001) made an update of the first review and found 14 articles containing data on the impact of trial entry on patient outcome. Eight of these 14 articles reported significantly better clinical outcome among trial patients than among non-trial patients. Three articles reported a favorable trend with trial participation. One article claimed to demonstrate an improvement in symptoms for patients in both the treatment and the control group. The two remaining articles found no evidence supporting any effect between trial participants and non-trial controls. Overall, Braunholtz et al. (2001) concluded that "there is little good quality evidence available" (p. 223), and added that, if anything, randomized trials have a beneficial rather than a harmful effect on the outcome of the participating patients. Finally, the authors pointed out that the best available evidence comes from clinical trials in oncology and conclusions should probably be restricted to that area only. The Emergency Care Research Institute (ECRI Institute 2002) found 10 comparisons and stated in their review that "some evidence shows that patients in phase II/III trials survive longer than similar patients who are not in trials" (p. 36). However, the authors noticed that only five of these comparisons controlled for differences in patient characteristics. Despite that in four of these five comparisons patients in trials survived significantly longer than those not in trials, the overall conclusion was that "one cannot have confidence in these results, […], due to the small evidence base" (p. 36). Similarly, Peppercorn et al. (2004) concluded that the available data is insufficient for claiming a trial effect. The authors reviewed 26 comparisons of outcomes for cancer patients who participated in clinical trials with those who did not. Fourteen of these 26 comparisons provided evidence of better outcomes among patients enrolled in trials. However, only eight comparisons (out of the 26) restricted non-trial patients to eligible patients for the trials in question. Only three out of these eight comparisons showed better outcomes in trial participants. In addition, their observations suggested that positive findings were more likely to be found in older studies conducted before 1986 and that the methodologies in most studies were inconsistent and incomplete; in particular, multivariable analysis was inconsistently applied. Finally, the authors of

a later Cochrane review concluded that outcomes of patients participating in RCTs do not differ from those of patients receiving similar treatments but not participating in RCTs (Vist et al. 2008). This review included five RCTs (6 comparisons) and 80 cohort studies (130 comparisons), with 86,640 patients treated in RCTs and 57,205 patients treated outside of RCTs. Taken together, it seems difficult to establish a definite link between enrolment in clinical trials and improved clinical outcome. Even though there is no good evidence that patient outcomes are worsened by trial participation, the available data suggests that individual trial participation does not offer substantial benefits.

In most of the publications, the term "trial effect" is associated with the result from comparisons between "trial patients" and "non-trial patients" (i.e., Question 1). In fact, different "trial" groups and particularly various "non-trial" groups are possible (Table 1 in Braunholtz et al. 2001). The trial group, for example, may consist only of trial participants but it may also contain all patients of recruiting physicians. The corresponding non-trial group may include patients who refused enrollment into a trial or patients of recruiting health care providers but who were not invited for participation in the trial. Finally, eligible patients of non-recruiting health care providers or simply all their patients may serve as controls. In the last years another interesting aspect of participation in clinical trials has come up: Instead of comparing trial patients with non-trial patients, it has been argued that it may be even more meaningful to compare institutions that participate in clinical trials to those that do not. The idea behind this institutional trial effect is that engagement in clinical research may improve health care performance through various mechanisms and that institutions might improve the health outcomes of their patients just by becoming research active. Consequently, Question 2 has been formulated as: "Do health care institutions or service providers who are active in research deliver better care and outcomes than those who do not participate in clinical research?" (Selby and Autier 2011, p. 6). In contrast to individual trial participation, up to now, only a few studies have explicitly explored the relationship between institutional research activity and patient outcomes. Indeed, the reviews mentioned above either focused on the comparisons between trial patients and non-trial patients or did not really distinguish between the different levels of trial participation and non-participation, and thus did not differentiate between the two questions cited above.

One review published by Clarke and Loudon (2011) searched the literature for studies directly addressing the effects of research activity at the structural level (Question 2). The authors found five articles that examined the impact of trial participation of health care practitioners and eight articles that assessed the effect of institutional participation in clinical trials on patient outcomes. They judged the studies as "controlled" when the analyses were adjusted for potential confounders. When selection bias could not be entirely excluded because of possible differences in patient mix, the studies were labelled as "poorly controlled". Overall, Clarke and Loudon (2011) concluded that there might be a beneficial effect of institutional research activity on outcomes but stated that "[…] the consequences for patient health are uncertain and the most robust conclusion may be that there is no apparent evidence that patients treated by practitioners or in institutions that take part in trials do worse than those treated elsewhere" (p. 1). Moreover, the authors acknowledged that the available findings were inconsistent and difficult to integrate. The heterogeneity of the studies made a quantitative meta-analysis impossible. Therefore, the studies on the trial effects related to being treated in research-active institutions will be presented separately in the following paragraphs.

The three studies that were judged as “poorly controlled” suggested positive effects of research activity. Clark et al. (2003) studied data on aphaeresis use in 19 institutions active in trials and five non-active institutions, and found higher aphaeresis use in trial centers compared to non-trial centers. However, the aphaeresis use increased in both types of institutions during the observation period. Janni et al. (2006) investigated centers before and after participation in a German trial of chemotherapy regimens for women with breast cancer. In this study, trial participation led to an increase of the relevant information flow through regular newsletters and study meetings and this improved professional knowledge in 80% of hospitals; in 31% of the centers the delivered care got better. According to the authors “these results support the hypothesis that carrying out the study has a positive effect on the current medical care of participating patients, irrespective of the knowledge gained later from the actual findings of the study” (p. 3665). Chen et al. (2006) studied the impact of hospital type on treatment and found that the uptake of changes in practice for the treatment of laryngeal cancer was greater in teaching/research facilities than in community hospitals and community cancer centers. The authors speculated that the reasons for lower uptake of new treatments at community institutions compared to institutions involved in teaching and research activities “may be lower awareness of treatment advances, lack of multidisciplinary expertise or availability of specific treatments at the facility, or referral bias” (p. 837).

Five studies on institutional research activity were assessed as “controlled” but these studies provided mixed results. Majumdar et al. (2008) looked at patients with coronary artery syndrome treated at hospitals with no, low or high trial participation, and found that in-hospital mortality significantly decreased with increasing level of trial participation: 5.9% vs. 4.4% vs. 3.5%, respectively. Moreover, compared with hospitals that did not participate in trials, those hospitals that were active in trials had higher adherence to treatment guidelines and seemed to provide better care. This finding suggested that the delivery of health care may also be affected by trial participation of the institution, and that this may have a positive impact on all patients accessing the service, not just on those actually participating in trials. In contrast, an earlier study by Majumdar et al. (2002) investigated the prescribing practice following myocardial infarction in North America. This study did not find any significant differences in treatment of patients between trial and non-trial hospitals. This study anticipated that hospitals that had taken part in the Survival and Ventricular Enlargement (SAVE) trial would become early adopters and treat their patients with angiotensin-converting enzyme (ACE) inhibitors that were proved to be effective following myocardial infarction. However, even after the study closed, patients treated at hospitals that had taken part in SAVE were not more likely to receive an ACE inhibitor than were patients treated at non-SAVE sites (16% vs. 15%). Similarly, a Danish study comparing 10 practices participating in a trial of an asthma treatment with 165 non-trial control practices did not detect any differences in physicians' adherence to treatment guidelines. Interestingly, trial participation seemed to affect physicians' prescription practice in that way that physicians involved in the trial conduct were more likely to prescribe medication sponsored by the pharmaceutical company (Andersen et al. 2006). A Canadian study in patients with myocardial infarction found that trial participants had better survival than patients treated in the same hospitals but not actively treated in trials or those in non-trial hospitals (Jha et al. 1996). The last of the “controlled” institutional studies included in the review found higher survival rates for ovarian cancer patients and better adherence to treatment guidelines in trial hospitals compared to non-trial hospitals (du Bois et al. 2005a). The data of this German study is part of the following analyses, and will be, therefore, described in more detail later in this thesis.

Later reviews, published by Hanney et al. (2013) and by Boaz et al. (2015), were the most comprehensive and focused on mechanisms through which research activity might improve health services performance and patient outcomes. One part of these reviews dealt with the impact on health care that had arisen as so-called byproduct of the research involvement in the original study. This part consisted of a total of 21 papers, including 12 of the papers from Clarke and Loudon's (2011) review. In this byproduct category of papers, the main purpose of the original research activity of clinicians or organizations was to conduct or participate in studies to evaluate the efficacy of new therapies or procedures. Of the 21 identified papers, 17 papers produced positive findings and four papers produced negative or mixed results about the association between engagement in research and health services performance. The authors concluded that the current evidence does suggest that a link exists. However, they acknowledged that this is more likely to be demonstrated through improved health care processes than through improved patient outcomes per se. Finally, the 21 papers in the review were judged by the authors to be even more diverse than the 13 papers included in the previous review performed by Clarke and Loudon (2011), so that a quantitative meta-analysis was not conducted.

It is obvious from above that the literature on trial effects has lacked clarity. First of all, there is a need for careful formulation of questions that address the effects of participation in clinical research. In particular, possible effects of research activity on the outcomes of individual patients involved in clinical trials (Question 1) have to be distinguished from the impact of research activity on the outcomes of health care institutions (Question 2). The participants of the IARC workshop agreed that the first question has now been extensively investigated without coming to any convincing conclusions (Selby 2011). They, therefore, accepted the current evidence from literature showing that there is probably little or no impact on individuals entering clinical trials compared to similar patients treated in similar institutions but outside trials. In contrast, questions relating outcomes of whole services to their research activity have been less well defined and have been much less investigated, despite their probably even higher relevance to health care providers, policy makers, and patients. Consequently, even though the widely-held opinion is that institutional engagement in clinical research is beneficial, the evidence that research activity improves health care performance and outcome is less strong than expected. In line with this, the participants of the IARC workshop judged the available data to be too limited to make definite recommendations. Nevertheless, Question 2 was considered crucial and it was agreed that further investigation of this topic is required to determine whether clinical research benefits patients and health care systems within which it is done. The main reason for this conclusion was that Question 2 is relevant to all patients in a health care system, not just those patients who are enrolled in clinical trials: If institutional research activity has positive effects, these benefits affect many more patients than just the small proportion of actual trial participants. In addition, with the widespread recognition that outcomes may vary by type of provider, health institutions are increasingly being asked by policy makers to report data on the quality of care they deliver. This includes results on possible relationships between trial participation of institutions and the outcomes of their non-trial patients (see also Clarke and Loudon 2011). Finally, Question 2 was not only considered central to health care policy but also to the patients—the actual users of health care services. Nowadays, patients check multiple sources of information when making treatment choices. Consequently, the potential effects of research activity are also important to patients who wish to use the available evidence when deciding about where they receive their health care and from whom.

As a starting point to address Question 2 it is important to understand of what is meant by quality of health care.

1.2 Quality of health care

Quality of health care can be popularly described by the simple premise of "doing the right things right" (Birkmeyer et al. 2004, p. 632). While there are many similar scientific definitions, the Institute of Medicine (IOM) of the National Academy of Sciences in the United States has proposed one that is widely accepted and covers the main features of many others. The IOM defines quality of care as "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" (Lohr 1990, p. 4). This definition, however, does not provide guidance on how to measure quality of care. In practice, several approaches focus on different measures used to reflect quality in a given system. One of the most prominent is a framework for evaluating quality of care proposed by Avedis Donabedian more than 50 years ago. Donabedian (1966) suggested assessing quality of care within a triad comprising of structure, process, and outcome. The model is most often illustrated by a chain of three boxes, representing different kinds of information that may be collected to draw conclusions about quality of care. Unidirectional arrows, meaning that each component has a direct influence on the next, connect structure, process and outcome (see Figure 1). In the following, these components will be described in detail, and potential advantages and limitations of their use will be highlighted.

Figure 1: Donabedian's (1966) concept of health care quality

In the terminology of Donabedian (1966; 1980; 1988), *structure* includes all the factors that affect the context in which care is given to patients. Structure refers to attributes of the personnel who provide care and the setting in which care is delivered. Human resources, for instance, include the number and the qualifications of staff such as education, training, job specific skills, and professional experience of physicians and nurses. Material resources consist of facilities such as buildings, equipment, and technology but also resources in terms of money and available time. Finally, structure denotes the attributes of organizational infrastructure such as administration, methods of peer review, and organizational incentives. Examples for structural attributes specific to oncology are the physician's specialty and board certification or the institutional specialization and annual procedure volume. Especially the latter has been frequently used as a measure for the structural aspect of quality, and volume-outcome relationships have been described for various types of surgical procedures (Hillner et al. 2000). Even though the respective literature has not been consistent (Gruen et al. 2009), high-volume providers were reported to have lower operative mortality, fewer complications, and better long-term survival than their lower-volume counterparts (Birkmeyer et al. 2004). Another structural variable often cited as a predictor of improved surgical outcomes in oncology is subspecialty training of the operating surgeon (Earle et al. 2006; du Bois et al. 2009b). Thus, patients are often encouraged to seek care at hospitals with high volumes for several procedures and centers with specific expertise.

A rationale for focusing on structure is the premise that, given the right setting, high-quality care will follow—"good structure increases the likelihood of good process" (Donabedian 1988, p. 1747). For example, one would expect care to be of high quality when all staff members have some level of expertise in the field of interest, know their roles and responsibilities, are well coordinated, and have implemented strategies for continuously improving quality of care. In contrast, if the institution is an unpleasant place to work, roles and responsibilities are not clear, decision-making processes are complicated, and the resources are restricted due to budget or time constraints, the employees will probably not do a good job. In addition, structure is increasingly viewed as not just the way health care providers and institutions are organized and operated, but also by the policies they have in place and that affect quality of care. For example, processes for monitoring and auditing quality and promoting high-quality care can have impact on how well care is eventually delivered. Recent examples of initiatives related to these structural aspects include certification systems such as the German OnkoZert (https://www.onkozert.de/). This certification confirms that processes as well as technical and scientific standards of oncology care are in accordance with the requirements of the German Cancer Society, thereby ensuring patients receive optimal cancer treatment.

Birkmeyer et al. (2004) suggested that structural variables have been successful mainly due to their convenience, as they are tangible, easily assessed, and inexpensive, often utilizing administrative data. It is relatively simple, for example, to determine whether a hospital is a teaching or a non-teaching hospital and whether it specializes in a particular type of care. Similarly, it is quite easy to figure out whether incentives are provided to doctors who meet high quality standards. However, even though structure is relatively straightforward to assess, the links between structure and process variables or even outcomes are quite complex (Landon et al. 1998) and, as a consequence, often weak (Landon et al. 2001). This is indeed the main disadvantage of structure measures, and most studies have had difficulty in linking structural variables consistently to process measures and outcomes at patient level. Among other reasons, this is because the most available attributes of infrastructure often lack specificity. Thus, only little is known about the relevance of structural characteristics that are more challenging to determine. For example, one can easily find out if a physician is certified by an approved board in oncology, but it is a great deal to measure the quality of that certification and it is even harder to determine to which extent the physician uses the acquired knowledge or specific skills in daily routine. Similarly, it is rather straightforward to find out if a hospital is affiliated with a medical school and provides clinical education and training to health professionals, but it is much more difficult to assess the actual teaching quality. According to Birkmeyer et al. (2004), the literature assessing structure is, therefore, incomplete because it focuses on a very small number of key measures. In addition, unlike process measures, most structural variables can only be examined in observational studies. Thus, it is much more difficult to rule out confounding as a possible explanation for observed effects. Moreover, structural variables mostly only reflect average results for groups of health care providers, and not for individuals. Finally, many structural variables are not actionable from a health care provider's perspective, which limits their effective use as a means towards quality improvement. For example, a low-volume hospital cannot easily make itself a high-volume hospital for a specific service and a hospital without appropriate infrastructure cannot simply become a specialized center. Similarly, a general gynecologist or surgeon cannot just become a specialist in gynecologic oncology overnight. Changes like this need serious commitment, investments and resources: education, equipment, and especially time. All of the above make structure variables only imperfect surrogates of quality.

Process variables describe the health care that patients actually get. In Donabedian's (1966) terms, process refers to all the activities of providing care that use resources and produce outcomes. These activities often include diagnosis, treatment, and preventive actions but may be extended to efforts undertaken by the patients themselves or their relatives to improve health outcomes. A common example in oncology is administration of adequate chemotherapy to patients with advanced stage of disease in a timely manner. Pain self-management in oncology is another example on the patient side. Processes can also include physician-patient interactions such as providing emotional support and information, as well as involving patients in decision-making in a way that fits their preferences. Process measurement can include both: what is done (i.e., whether the right actions were undertaken) and how it is done (i.e., how well actions were carried out). In other words, process measures reflect whether a patient received what is generally recognized to be the right care. According to Donabedian (1966; 1980; 1988), the measurement of process variables is thus nearly identical to the measurement of quality of care as process contains all actions of care delivery in clinical practice. Still, process variables can be easier to measure than outcomes because information about them can be obtained, without delay, from medical records, surveys or interviews with patients and physicians, or even from direct monitoring of care services.

A potential limitation of process measures is the lack of data that confirm their value for specific aspects of care. The consequence is that much care is delivered in the absence of convincing evidence for its effectiveness (e.g., the value of guideline-based treatment for elderly patients is often not definitively known). Another big challenge about process measures is related to the identification of exactly those processes that are linked to improved outcomes. Especially the identification of valid quality metrics for complex procedures, as well as the definition of methods to measure and report them, remain a challenge. Process measures are typically defined based on the known link between a certain aspect of treatment and a certain outcome. However, even when data is available that supports the effectiveness and appropriateness of a process, there is often more than one evidence-based option to treat a certain condition. Sometimes there are dozens of things that can be done in the course of taking care of a patient with a complicated disease and little information is available about which processes are crucial for achieving the best results for the patient. Consequently, it is increasingly recognized that simple assessment of individual processes of care is probably not adequate. Instead, bundles of multiple tasks need to be done to achieve the best possible outcome. Another practical disadvantage of process measurement is that one must be able to accurately identify eligible patient populations. Many processes known to be generally effective may not be appropriate or may not be even desirable for individual patients (e.g., chemotherapy treatment with all its side effects). Further limitation of process measures is that sometimes it is not known how certain processes of care are related to outcomes. On the plus side, process variables have some attractive features. Measures of process of care are generally sensitive indicators of quality of care, and they may explain apparent links between structure and outcomes. Moreover, process variables are very specific and may thus be perceived by health care providers as "fairer" measures of quality than structural measures (Birkmeyer et al. 2004). In particular, they answer the straightforward question doctors and nurses might have: "Am I doing the right thing for my patient?" Consequently, they are explicit and intuitive to understand for health care providers and often for patients too. Finally, and most importantly, they are actionable from the practitioner's perspective because they clearly indicate what is expected and what should be done how, and what might be changed to improve.

Quality of care can be assessed in terms of direct *outcome* measures, which seek to capture whether the goals of health care were achieved. They can be measured at hospital as well as on individual patient level. For example, hospital standardized mortality ratios which compare actual to expected in-hospital deaths can help identify potential areas of improvement and facilitate benchmarking against other institutions. In practice, health care providers use patient outcomes to track, monitor and evaluate changes in patient health but also to guide their treatment. Direct outcome measures are often seen as the most important indicators of quality because the primary goal of health care is to improve patients' health status. Health care providers are most familiar with outcome components that describe the end goal of quality care (e.g., survival associated with certain surgical procedures). However, such measures sometimes do not adequately reflect the possible range of health conditions that are affected by care and that are important to individual patients. Other outcomes may be health-related quality of life, patient satisfaction, complication or readmission rates, or even financial costs.

The main advantage of direct outcome measures is their face validity. In the quality of care arena, improved survival or improved health status are most desired by patients and are, thus, subjectively viewed as covering the important results of health care. In addition, they are likely to get the greatest buy-in from health care providers. However, accurately measuring outcomes and relating them exclusively to delivered health care might be difficult. Although improving outcomes is clearly the goal, assessing outcomes may be difficult because of the long-time frame needed to achieve improvement and other potential factors besides quality that may affect outcomes. If outcomes are considered to serve as quality measures they must reflect or be at least responsive to changes in health care. In addition, they may need substantial time to become visible, so that often long-term follow-up is needed. A further potential limitation is that even if outcomes are not optimal, they may not provide insights into why they were suboptimal and what needs to be changed to make them better. Another practical difficulty in assessing outcomes is that outcome data may not be routinely available in clinical practice. Moreover, for some patient reported measures, selection bias can occur if patterns of responders differ between settings because patients with worse outcomes may be less likely to return a self-reported questionnaire. Finally, the most important limitation of outcome measurement relates to sample size. For the majority of procedures, very few health care providers have sufficient events and cases that would allow for meaningful, procedure-specific measures of morbidity and mortality (Birkmeyer et al. 2004). On the other hand, large sample sizes and adjustments for case mix are required for drawing valid conclusions about relationship between processes and outcomes.

To sum up, Donabedian's (1966; 1980; 1988) model has been widely accepted as a useful framework for assessing quality of care in clinical practice. However, each of the three components has its unique advantages and limitations. Thus, the assessment of quality and the corresponding efforts towards improvement should ideally incorporate all three dimensions of care, which is sometimes difficult to realize. In the quality-improvement literature, structural measures, for example, have received considerably less attention than processes and outcome measures. According to Birkmeyer et al. (2004), both health care providers and policy makers should thus be flexible in their approaches to measuring quality of health care and align on strategies best suited to meeting their specific goals. Recommendations for quality assurance and improvement using outcome measures should be selected and prioritized on the basis of their actionability and validity. Most importantly, the metrics assessed should offer valuable insights and enhance the quality of care delivered to patients.

1.3 Ovarian cancer

Cancer is one of the leading causes of death worldwide and despite better treatment options, it is still a crucial obstacle to enhancing life expectancy in most countries around the globe (Bray et al. 2021). According to the Global Cancer Observatory (GLOBOCAN) database provided by IARC, 19 million cancer cases were newly diagnosed and almost 10 million people died from cancer in 2020. The global cancer burden is expected to be 28 million new cases in 2040, a 47% increase from 2020 (Sung et al. 2021).

Ovarian cancer ranks eighth in incidence and mortality among women, with 313,959 diagnoses and 207,252 deaths in 2020 (Ferlay et al. 2020a). Ovarian cancer incidence rates have been recently declining or stabilizing (Coburn et al. 2017; Cabasag et al. 2020a), but a woman's lifetime risk remains 1 in 78 for developing and 1 in 108 for dying from ovarian cancer (American Cancer Society 2023). Ovarian cancer prognosis is still poor due to delayed diagnosis and unequal care access, resulting in the highest mortality rate among gynecological cancers (Cabasag et al. 2020b). By 2040, a 42% increase in cases (445,721) and a 51% increase in deaths (313,617) are projected by GLOBOCAN (Ferlay et al. 2020b). Some incidence and mortality rates for ovarian cancer in 2020 are shown in Table 1. They are presented as crude numbers as well as age-standardized rates (ASR) per 100,000, which can be used for comparisons between regions. Regional variations exist, with Central and Eastern Europe having the highest incidence and Africa the lowest. Mortality rates have declined in high-income compared to low-income parts of the world over the last two decades, but these trends are less pronounced than those for incidence (Malvezzi et al. 2016; Dalmartello et al. 2022).

In Germany, 7,162 diagnoses and 5,326 deaths occurred in 2020, with an incidence ASR of 7.3 and a mortality ASR of 4.4, both rates being slightly higher than the respective rates for Western Europe (Ferlay et al. 2020a). By 2040, 7,860 new ovarian cancer diagnoses and 6,197 deaths from ovarian cancer are expected in Germany (Ferlay et al. 2020b).

Table 1: Ovarian cancer incidence and mortality age-standardized rates by selected world regions*

	Incidence		Mortality	
Region	No. of cases	ASR/100,000	No. of deaths	ASR/100,000
World	313,959	6.6	207,252	4.2
Central-Eastern Europe	28,530	10.7	17,565	5.6
Northern Europe	9,457	8.8	6,530	5.0
Southern Europe	12,779	8.0	8,015	4.0
Western Europe	15,927	7.1	11,943	4.3
Northern America	26,630	8.1	16,451	4.1
Central America	5,987	6.1	3,698	3.7
South America	16,176	5.8	10,579	3.6
Asia	170,759	6.2	112,936	4.0
Africa	24,263	5.4	17,008	4.0
Australia/New Zealand	1,717	6.4	1,265	3.9

Note: *from the GLOBOCAN 2020 database, ASR = Age-standardized rate (per 100,000 person-years)

1.3.1 Prognostic factors

Most ovarian cancers (approximately 90%) are epithelial carcinomas, meaning that the cancer begins in the cells on the surface of the ovary (Ledermann et al. 2013). The remaining ovarian tumors originate from the germ or stromal cells. Only about half of the women with epithelial ovarian cancer are still alive five years after diagnosis (National Cancer Institute 2023).

Over the last decades, researchers have developed several prognostic models and have identified multiple factors associated with prognosis of ovarian cancer. The most prominent proposals are statistical prognostic tools that provide a quantitative estimate of the survival probability for an individual patient with epithelial ovarian cancer. Two of them, both nomograms, were developed to predict survival for all stages of disease (Clark et al. 2001; Barlin et al. 2012). Three others, two nomograms (Chi et al. 2008; Gerestein et al. 2009) and one prognostic index (Teramukai et al. 2007), were developed exclusively for patients with advanced-stage disease. Van de Laar et al. (2014) externally validated the latter three prognostic models for survival in patients with advanced-stage ovarian cancer. Clark et al. (2007) also externally validated one of the models (Teramukai et al. 2007). All models suggest that the most important unfavorable patient and disease characteristics in ovarian cancer are higher stage of disease, older patient age at diagnosis, bad performance status, poorly differentiated tumor, clear-cell and mucinous histology, and presence of ascites. The following sections therefore briefly describe these risk factors and summarize their importance in the prognosis of ovarian cancer patients.

The extent of the disease is generally expressed in terms of its *stage* (Pecorelli et al. 1999). The staging of gynecological cancers has been standardized by the International Federation of Gynecology and Obstetrics (Fédération Internationale de Gynécologie et d'Obstétrique, FIGO). According to FIGO, ovarian cancer staging should be done surgically. The tumor stage at diagnosis is considered to be the most important prognostic factor for survival of patients with ovarian cancer (Heintz et al. 2006), and in their review, Clark et al. (2001) identified stage of disease as one of the most consistently reported prognostic factors. Women diagnosed with early-stage (FIGO I–IIA) ovarian cancer have a more favorable prognosis than women diagnosed with more advanced stages (FIGO IIB–IV). When ovarian cancer is detected very early (FIGO stage I), meaning that the cancer is still limited to the ovaries at diagnosis, about 80–90% of women survive longer than five years (Heintz et al. 2006). Unfortunately, ovarian cancer lacks of early and specific symptoms and has thus been called a “silent killer” (e.g., Jasen 2009). As a consequence, most patients are not diagnosed until the disease has advanced and this usually means that the cancer has already spread beyond the ovaries (Pecorelli et al. 1999). FIGO stage II ovarian cancer is a small and heterogeneous group that is defined as extension or metastasis to nearby regions in the pelvis, most commonly the fallopian tubes and pelvic peritoneum. This group comprises only about 10% of ovarian cancers and has still a 5-year survival rate of about 70%. Most commonly, ovarian cancer presents in FIGO stage III, comprising about 50% of cases. Stage III tumors characteristically spread along peritoneal surfaces involving both the pelvic and the abdominal peritoneum. Heintz et al. (2006) reported 5-year survival rates of 47% among patients with FIGO stage IIIA, 42% with FIGO stage IIIB, and 33% with FIGO stage IIIC. About 13% of patients present in FIGO stage IV disease. This stage includes patients with tumor growth involving one or both ovaries with distant metastases. The 5-year survival is only 19% for women with FIGO stage IV disease (Heintz et al. 2006). Despite awareness of the disease, survival trends have not really changed since then. The main challenge is still the early diagnosis of ovarian cancer. Hence, recognizing symptoms with suitable diagnostic tests continue to be crucial in enhancing patient outcomes.

Age at time of diagnosis is considered a further important prognostic factor for patients with ovarian cancer (Thigpen et al. 1993). Epithelial ovarian cancer is predominantly a disease of postmenopausal women, with the majority (> 80%) of ovarian cancers occurring after the age of 50 (Ledermann et al. 2013). The incidence of ovarian cancer increases with age and has its peak in the eighth decade of life. The median age of diagnosis is 63 years, and 48% of patients are 65 years or older (Ozols et al. 2006). Tumors in younger women have a more favorable prognosis than tumors in older women. Five-year survival for women younger than 65 years of age is about 66% compared to 33% for women 65 years and older. Age is assumed to affect prognosis in ovarian cancer in two ways. On the one hand, possibly due to different tumor biology in the aged, there seems to be an independent association of advanced age with prognosis. On the other hand, it has been speculated that the management of ovarian cancer in elderly women is more conservative; meaning that exposure to more aggressive treatment is lower than in younger women (Yancik 1993).

Performance status is a rather crude assessment of how well a patient is able to perform ordinary tasks and to carry out activities of daily life. The Eastern Cooperative Oncology Group (ECOG) scale is widely used by practicing oncologists and researchers to quantify the functional status of cancer patients. The ECOG scale describes the status of symptoms and functions with respect to ambulatory status and need for care (Sørensen et al. 1993). ECOG performance status 0 means normal activity, ECOG 1 is associated with some symptoms, but still near fully ambulatory, ECOG 2 means less than 50%, and ECOG 3 means more than 50% of daytime in bed, while ECOG 4 is associated with complete disability. ECOG 5 corresponds to patient's death. Although performance status suffers from problems of subjectivity it has been demonstrated to be an independent prognostic factor for survival in ovarian cancer (Lund et al. 1990). Moreover, it has been shown that patients with a good performance status (ECOG 0 or 1) are more likely to respond to treatment and experience fewer and less severe side effects from treatment (Ozols et al. 2006).

Tumors are *graded* according to how well or poorly differentiated they are. Worldwide, there are many different systems for the grading of ovarian cancers. Some of them are derived from reviewing the following tumor characteristics: architectural features, mitotic counts and nuclear atypia. Based on these characteristics, most proposed grading systems are systems with three grades (Shimizu et al. 1998): According to the World Health Organization (WHO) epithelial ovarian cancers are graded as either well differentiated (Grade 1, G1), moderately differentiated (Grade 2, G2), or poorly differentiated (Grade 3, G3), with a further category that applies if grade cannot be assessed (Grade X, GX). In some systems, Grade 4 (G4) is used to describe undifferentiated high grade tumors. While most of the early stage cancers are well differentiated, poorly differentiated tumors are most common in the advanced stages. However, there is no single universally accepted grading system and ovarian carcinomas can also be classified into either low grade or high grade tumors (Ledermann et al. 2013). In general, tumor grade can serve as an additional prognostic factor in ovarian cancer. Low-grade tumors have been reported to be associated with a more favorable prognosis than high-grade tumors. Patients with FIGO stage I and poorly differentiated tumors have worse survival compared with patients at the same stage but with well-differentiated tumors. The 5-year survival rates for women with early-stage ovarian cancer (FIGO I and II) are 90%, 80%, and 75% for G1, G2, and G3, respectively. For advanced ovarian cancer (FIGO III and IV), the reported 5-year survival rates are 57%, 31%, and 28%, respectively (Heintz et al. 2006).

The WHO *histological classification* of epithelial ovarian tumors recognizes various distinct subtypes such as serous, mucinous, endometrioid, clear cell, mixed, undifferentiated, and unclassified tumors (e.g., Kaku et al. 2003). Invasive serous carcinomas are the most common histological subtype accounting for up to approximately 80% of advanced ovarian cancers. In contrast, clear-cell cancers account for approximately only 5% of ovarian carcinomas. Mucinous carcinomas are rather uncommon histological types that affect several organ sites (Kelemen and Köbel 2011). Some studies have shown that the histological subtype can be important for ovarian cancer prognosis (Ledermann et al. 2013). For example, Winter et al. (2007) reported clear-cell histology to be associated with a worse overall survival compared with serous carcinomas. In advanced-stage disease, du Bois et al. (2009a) showed that compared with the most common serous histological subtype the mucinous histological subtype predicted inferior survival outcome. Similarly, Mackay et al. (2010) have shown that mucinous and clear-cell carcinomas were independent predictors of poor prognosis in patients with advanced stages of ovarian cancer. Nevertheless, the histological subtype has generally been considered to have less prognostic value than other clinical factors such as stage of disease (Clark et al. 2001; Ozols et al. 2006). In their review, Clark et al. (2001) stated that histology and grade were both included in most prognostic models, but were not consistently statistically significant.

Ovarian cancer is the most common primary cancer site associated with *ascites* (Shen-Gunther and Mannel 2002; Kipps et al. 2013). The incidence of ascites in women with epithelial ovarian cancer ranges from 45% to 75% depending on the tumor type and increases in advanced stages (Partridge and Barnes 1999). Presence of ascites is not only one of the symptoms of ovarian cancer but it is, in general, also considered a poor prognostic indicator for survival. According to Puls et al. (1996), development of ascites correlates with a significantly decreased 5-year survival rate among women with FIGO stage III or IV epithelial ovarian cancer (5% with ascites versus 45% without ascites). Similarly, ascites has been found to be a poor prognostic factor in women with FIGO I disease (Dembo et al. 1990).

1.3.2 Treatment options

Several guidelines and consensus statements regarding surgical procedures as well as chemotherapy treatment have been published worldwide for both early and advanced ovarian cancer over the past decades (e.g., NIH Consensus Conference 1995; European Society for Medical Oncology (ESMO) Guidelines Task Force 2001; see Karam et al. 2017 and Vergote et al. 2022 for the Fifth and Sixth Ovarian Cancer Consensus Conference Statements). Surgery is the cornerstone of management of ovarian cancer. It is not only a critical part of effective treatment, but it is also required for diagnosis and staging (Pecorelli et al. 1999). Staging is the assessment of how far the tumor has spread. Proper and accurate staging conducted via surgery can separate patients with true early-stage disease from patients with advanced stages of ovarian cancer. This distinction is essential in guiding further treatment decisions, particularly when it comes to determining the necessity of adjuvant chemotherapy. A few early-stage cancers are curable with surgery alone (Heintz et al. 2006) but most women with ovarian carcinoma need chemotherapy in addition to surgery. Incomplete staging can lead to underestimation of the extent of the disease and thus result in undertreatment that may in turn affect patient outcome. The German S3-Guideline on diagnostics, therapy and follow-up of malignant ovarian tumors (Wagner et al. 2013; Staebler and Mayr 2017) still recommends the following surgical steps: longitudinal laparotomy, inspection and palpation of the entire abdominal cavity, peritoneal cytology, biopsies from all abnormal sites and peritoneal biopsies from unremarkable regions, bilateral excision of adnexa of uterus, hysterectomy, infracolic omentectomy, appendectomy, bilateral pelvic and paraaortal lymphonodectomy.

Surgical outcome in ovarian cancer is usually classified according to the amount of postoperative residual tumor. Resection is regarded as complete if no macroscopically visible tumor is left. If tumor remains after surgery, it is classified according to its largest diameter. In a landmark study, Griffiths (1975) showed, for the first time, an inverse relationship between residual tumor diameter and survival. Since this study, postoperative residual tumor has been consistently reported as one of the most powerful determinants of survival in ovarian cancer patients (e.g., Hoskins et al. 1992; Clark et al. 2001). However, the definition of "optimal" debulking surgery has varied over time and the cut-off for determining optimal cytoreduction has been changed repeatedly. Originally, residuals up to 15 mm largest diameter were classified as optimal surgical outcome whereas any larger residual tumor was defined as suboptimal debulking (Griffiths 1975). In contrast, some other studies that sought to predict the surgical resectability of ovarian cancer, defined optimal cytoreduction as a residual disease measuring less than 20 mm largest diameter (e.g., Meyer et al. 1995). Finally, a residual tumor size of less than 10 mm has been generally accepted as the definition of optimal surgical cytoreduction since the work of Bristow et al. (2002) who showed in a meta-analysis a significant positive correlation between percentage of maximal cytoreduction and median survival after controlling for other variables. However, newer publications claimed that not 10 mm but only complete debulking to no macroscopic residual tumor should be the goal of the surgical management of ovarian cancer. For example, du Bois et al. (2009a) conducted an exploratory analysis of three prospective randomized trials that examined platinum-taxane based chemotherapy regimens for advanced ovarian cancer between 1995 and 2002. In this analysis, they categorized patients based on their surgical outcomes: complete resection, small residual tumor burden (1-10 mm), and macroscopic residual disease (>10 cm). Findings revealed that patients with complete resection had significantly better survival rates compared to those patients with residual tumors. The prognostic impact of optimal debulking was smaller when compared to cases with macroscopic residual disease. In line with this, the current German S3-Guideline as well as the Fifth Ovarian Cancer Consensus Conference statements (Karam et al. 2017) define the goal of primary surgery as complete resection. Systemic chemotherapy for ovarian cancer is most often based on a platinum compound alone or in combination with a taxane (du Bois et al. 2005b; Karam et al. 2017; Vergote et al. 2022). Patients with early-stage disease (except for FIGO IA G1) should receive a platinum-based chemotherapy. According to the German S3-Guideline, patients with stage IA G1 ovarian cancer after complete operative staging must not receive adjuvant chemotherapy. In contrast, patients with stage IA G2, IB G1/2 can be offered platinum-based chemotherapy, and patients with FIGO IC or IA/B, G3 ovarian cancer must receive chemotherapy that should include carboplatin and consist of six cycles. Since 20 years, the recommended first-line chemotherapy for patients with advanced ovarian cancer (FIGO IIB–IV) is carboplatin and paclitaxel over a total of six cycles, with one cycle every three weeks (Ozols et al. 2003).

In summary, the standard treatment for most ovarian cancer patients involves accurate staging with maximal cytoreductive surgery followed by primary platinum-based chemotherapy. Surgical outcomes depend on tumor resectability and patient condition but can be improved by skilled surgeons in specialized centers (Bristow et al. 2002). In contrast, the chemotherapy outcomes are mainly influenced by chemosensitivity of the tumor, a factor hardly amenable to alteration. However, a hospital's infrastructure can have an impact on adherence to treatment guidelines in terms of chemotherapy administration. In this thesis, both treatment factors will be examined as potential mechanisms driving the effect of institutional research activity on patient survival, while accounting for relevant patient and disease characteristics.

1.4 Mediation analysis

Investigating causal relationships is not a novel approach (Wright 1934). Assessment of mediation has been a key tool of statistical analysis in some disciplines for many years. Since the publication of the seminal paper by Baron and Kenny (1986), testing for mediation has been an integral part of statistical analysis, foremost in psychology. In epidemiology, both theoretical and practical aspects of mediation have been considered since Robins and Greenland (1992). Finally, mediation models have been used for assessing causal relationships in medicine where the mediating mechanisms continue to remain one of the greatest puzzles (Ruesch 1961).

More generally, mediation analysis is a tool to identify and explain the mechanisms that underlie an observed relationship between an independent variable *A* and a dependent variable *Y* through the inclusion of a third variable, the so-called mediator *M* (Figure 2). This third variable is intermediate in the causal path from the independent to the dependent variable (MacKinnon 2007). A mediating variable represents asymmetric relations between variables. Mediation also implies a temporal relationship with *A* occurring before *M* and *M* occurring before *Y*. In a mediation model, the independent variable is assumed to (at least partially) cause the mediator which is then assumed to cause the dependent variable. A mediator shares similarities with a confounder variable because it is related to both *A* and *Y*. But in contrast to a confounder that is not on a causal path between *A* and *Y*, the mediator explains the relation between *A* and *Y* because it transmits the effect of *A* on *Y*. According to Baron and Kenny (1986), a mediator "represents the generative mechanism through which the focal independent variable is able to influence the dependent variable of interest" (p. 1173). Mediators are also called process variables, thereby referring to their function as variables that describe the process by which an independent variable affects a dependent variable (Judd and Kenny 1981). The importance of mediation as the totality of underlying processes is well argued in Hafeman and Schwartz (2009) where they demanded the opening of the "black box" when investigating exposure-disease relations. Surrogate endpoints or intermediate outcomes can equally be considered mediators because these variables represent proximal measures of a distal outcome (Prentice 1989).

The core element of mediation analysis is the estimation of two effects: The natural direct effect of an independent variable (or exposure) *A* on a dependent variable (or outcome) *Y* and the natural indirect effect acting through the mediator *M*. Consider, for example, a very simple path model with institutional characteristics, such as hospital trial participation as exposure *A*, surgical outcome as potential mediator *M*, and patient survival as the final outcome *Y*. In Figure 2, the exposure has a direct effect on the final outcome; despite the name "direct", this path summarizes all the unknown sources by which *A* has an influence on *Y*. In addition, the exposure affects the mediator *M*, which, in turn, has an effect on the final outcome. The total effect is then the aggregate of these two effects. Specifically, the *natural direct effect* is the effect one would observe on the outcome if the exposure could be changed without inducing a change in the mediator. The *natural indirect effect* is the effect on the outcome one would observe if the mediator could be changed as it would when the exposure was manipulated (without actually changing the exposure). In the terms of the example, a non-zero natural direct effect would be a survival benefit in patients treated at research-active hospitals compared to hospitals not engaged in research; such a benefit would be observable even if there was not any difference in the surgical outcome between trial and non-trial hospitals. In contrast, the natural indirect effect refers to the difference in survival as a consequence of the difference in the surgical outcome between research-active and research-inactive institutions.

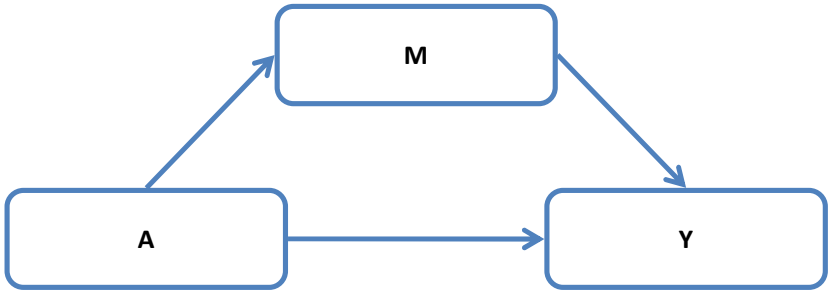

Figure 2: Simple path diagram relating exposure *A* to a mediator *M* and an outcome *Y*

For normally distributed mediators and outcomes, and in the absence of interactions or non-linear effects, natural direct and indirect effects can be estimated by a stepwise approach based on standard linear regressions. Baron and Kenny (1986) developed four steps in the evaluation of mediation by means of regression models. The first step is to show a significant overall association between exposure and outcome. Nonetheless, it's worth noting that the absence of a significant overall association does not preclude mediation, as the direct and indirect paths could counterbalance each other, resulting in a zero overall effect (Zhao et al. 2010). The next steps are to show that the indirect path exists, namely, that the exposure has an effect on the proposed mediator and that the proposed mediator has an effect on the outcome. The last step aims at showing that the relationship between exposure and outcome is numerically reduced after controlling for the potential mediator. This way, multiple regression models can be used for testing for direct and indirect effects. Models are fitted for the outcome both with and without the mediator, and the difference in the coefficients for the exposure is taken as a measure of the natural indirect or mediated effect.

However, researchers should be aware of the possible pitfalls associated with that traditional approach. In oncology, for example, the outcome of interest is often patient survival. Survival times are known to be non-normally distributed and typically right censored. Existing techniques for mediation analysis are thus not applicable to survival data, and several authors have shown that the linear regression approach cannot be simply translated to logistic regression or proportional hazard models (Cole and Hernán 2002; Kaufman et al. 2004; VanderWeele 2011). The main reason is that conditional effect estimates (e.g., the odds ratio from a logistic regression that includes the mediator) usually differ from marginal effect estimates (odds ratio from a logistic regression without the mediator), which suggests a mediator effect in Baron and Kenny's logic (1986). This may even occur in balanced designs and is, therefore, unrelated to the well-known Simpson's paradox.

For this reason, many studies in the field of health care investigate only the full regression models that simultaneously include structural variables as well as process measures (in the terminology of Donabedian 1966; 1980; 1988). However, because the process measures may be intermediate outcomes of structural measures, the effect of structural measures can be washed out or completely eliminated in such analyses. Appropriate modeling of structural and process variables needs to make a clear distinction between direct and indirect effects. A corresponding approach is implemented and discussed in this thesis (Lange et al. 2012; Lange et al. 2014; Rochon et al. 2014).

1.5 Objective and scope of the thesis

The general objective of the present thesis is to describe the quality of care for patients with ovarian cancer in Germany and to investigate the effects related to being treated in institutions that participate in clinical trials (Question 2, see Section 1.1 above). On the one hand, the thesis will focus on the relationship between institutional research activity, adherence to treatment guidelines, and patient outcomes. On the other hand, the thesis will deal with methodological challenges and will show possible solutions on how to investigate Question 2 in a proper way.

From the health care perspective, the thesis will examine whether institutional research activity can be considered as an indicator of high-quality care. The main hypothesis is that hospitals participating in clinical trials deliver better care and attain improved outcomes for their patients. For this purpose, data from three cohorts of a German quality assurance program in ovarian cancer will be explored with respect to adherence to treatment guidelines and with respect to survival in both early- and late-stage disease. One of the main purposes of the thesis will be then to proceed from the question "Does it work?" to the question "How does it work?" by investigating the relationship between measures of structure, process and outcome in one single Donabedian (1966; 1980; 1988)-like model. This is motivated by recent studies suggesting that patients treated in research-active institutions have better outcomes than patients treated in research-inactive institutions but generally paying little attention to the explanation of the observed effects. However, evaluating potential causal pathways for better outcomes in research-active institutions is critical to identify processes that should be targeted by quality improvement initiatives. Therefore, surgical outcome and chemotherapy administration will be investigated as potential mediators of the effect of hospital research activity on patient survival. The study results may help understanding gaps between clinical research and daily practice in oncology and prioritizing actions for the care of patients with ovarian cancer.

From the methodological point of view, the thesis will not just compare "on-trial" with "off-trial" patients, but all patients treated in hospitals participating in clinical trials with all patients treated in non-trial hospitals. In this way, the relationship between research activity (at the level of institutions) and outcomes (at the level of patients) will be explored while taking into account relevant patient and disease characteristics. The "How does it work?" question will be investigated by introducing a methodology that can be used to assess mediation on several types of mediators and outcomes including binary and survival data. In contrast to standard methods of regression analysis, this approach will allow for the first time to examine possible direct as well as indirect effects of institutional research activity on patient outcomes in one model and it will thus correspond directly to the conceptual model of structure, process and outcome cited above. In addition, the effect that a health care institution tends to attract similar patients and to provide care in a similar way will be explored, and it will be shown how potential clustering of patients within hospitals can be considered when doing mediation analysis. Finally, the methodological challenges of evaluating the impact of the clinical research process on institutions and health systems in general will be elaborated, and advantages as well as limitations of the conventional and the new methodological approaches will be discussed. Ultimately, this thesis will provide a clear recommendation on the most suitable methodology for future studies.

2 Materials and methods

2.1 Data

In Germany, a nationwide quality assurance program known as QS-OVAR (https://www.eierstock-krebs.de/qs-ovar/) monitors the treatment and outcomes in ovarian cancer. The Arbeitsgemeinschaft Gynaekologische Onkologie (AGO) Organkommission OVAR, a subcommittee of the German Cancer Society, initiated this program in 1999. Its aim is to describe the pattern and quality of care of patients with ovarian cancer as well as to improve their outcomes. The program was piloted in 2000 (du Bois et al. 2001a; b) but survival data was not collected at that point of time. Further data collections took place in QS-OVAR 2001, 2004, and 2008 with focus on quality of primary surgical and chemotherapy treatment while special attention was given to institutional characteristics. In addition, patient survival was collected for at least three years after diagnosis. These three cohorts are the subject of the present thesis. Du Bois et al. (2005a) described study design and first results from the cohort of patients diagnosed in the third quarter of 2001. Further details are provided in Rochon and du Bois (2011). QS-OVAR 2004 and 2008 were designed and conducted in a similar manner.

The study consisted of two phases. In Phase 1, all gynecological departments in Germany were invited to participate in a survey. They were requested to provide the number of patients newly diagnosed with ovarian cancer in the respective year, along with information about their affiliation with the two German study groups: the Ovarian Cancer Study Group of the Arbeitsgemeinschaft Gynaekologische Onkologie (AGO-OVAR) and the Northeastern Society of Gynecologic Oncology (NOGGO). Phase 2 involved the hospitals that responded to the survey. Approximately one year after the initial diagnosis of ovarian cancer, these hospitals were asked to document all patients newly diagnosed with ovarian cancer in the third quarter of 2001, 2004, or 2008. The study only included patients aged 18 or above who had histologically confirmed epithelial invasive ovarian cancer. Data were gathered using paper report forms, which were then cross verified with surgical and pathological reports before being entered into an electronic SPSS database (IBM Corp 2010). Medizinische Marktforschung (MMF GmbH, Dortmund, Germany) managed the data and queries. Trained data managers ensured the accuracy and plausibility of the data entry. All data were anonymized prior to analysis.

With regard to structural measures, data were collected on hospital volume, hospital care level, and hospital participation in prospective clinical trials conducted by at least one of the two study groups, AGO-OVAR or NOGGO. These two groups organize almost all national and international cooperative clinical trials for ovarian cancer in Germany. AGO-OVAR is a member of the Gynecologic Cancer Intergroup (GCIG), and both AGO-OVAR and NOGGO are members of the European Network of Gynecologic Oncology Trial Groups (ENGOT). Both study groups checked all hospital declarations with respect to their research activity. Some of the hospitals participated only in selected trials; some participated in trials of both study groups. However, the effect of participation in trials of the cooperative study groups was deemed to cover more than recruiting of individual patients in the third quarter of 2001, 2004 or 2008. To be considered as research-active, hospitals had to enroll patients in clinical trials of the two study groups before or during the third quarter of the respective year. Hospital volume was recorded as the number of patients who had been treated for ovarian cancer per year. Hospitals treating 1 to 11 patients annually were categorized as low-volume, and those treating 12 or more patients per year were categorized as high-volume hospitals. Finally, the hospitals were categorized into four categories with respect to their care level: university hospitals, hospitals with centralized services and maximum care, hospitals providing specialized care, and primary care hospitals.

For the purposes of this thesis, adherence to treatment guidelines for ovarian cancer was selected as the process measure of quality of care and was considered the therapeutic standard that the majority of ovarian cancer patients should be provided. This measure was based on national and international recommendations for surgery and chemotherapy according to the time period of diagnosis (Bauknecht et al. 2000; European Society for Medical Oncology (ESMO) Guidelines Task Force 2001; du Bois et al. 2005b). Adherence to treatment guidelines was defined as stage-appropriate surgery and chemotherapy in patients with early-stage (FIGO I–IIA) and advanced-stage (FIGO IIB–IV) disease. Dichotomous variables, adherence or non-adherence, were created with regard to surgical and chemotherapy guidelines separately. For the overall treatment (both surgery and chemotherapy taken together), categorical variables with four, three and two categories were built.

With respect to surgical staging in early ovarian cancer FIGO I–IIA, hospitals were asked to report whether the following nine procedures were done: 1) vertical laparotomy, 2) total abdominal hysterectomy, 3) bilateral salpingo-oophorectomy including 4) removal of all tumor tissue, 5) omentectomy, 6) peritoneal sampling, 7) cytology, as well as 8) pelvic, and 9) para-aortic lymph node staging. In patients younger than 50 years of age with highly differentiated FIGO IA tumors and an option for fertility-sparing surgery, total abdominal hysterectomy and bilateral salpingo-oophorectomy were not deemed mandatory. Staging was defined as “optimal” when maximally one step was missing and was considered “complete” when all surgical steps were performed. Adjuvant platinum-based chemotherapy was regarded standard care for patients with early ovarian cancer, except for patients with FIGO IA G1 tumors for whom a chemotherapy is not recommended.

For patients with ovarian cancer FIGO IIB and higher, treatment guidelines recommend surgery including maximal debulking as well as chemotherapy with carboplatin and paclitaxel (e.g., du Bois et al. 2005b). According to these guidelines, tumor residuals up to 10 mm largest diameter were considered as “optimal” surgical outcome whereas any larger residual tumor was defined as “suboptimal” surgical outcome. Because only complete debulking to no macroscopically visible residual tumor was recommended in more recent guidelines (e.g., du Bois et al. 2009a), “complete debulking” was assessed in sensitivity analyses. With respect to chemotherapy, any combination of platinum and taxane was considered adherent to treatment guidelines in patients with FIGO stage IIB–IV disease.

All patients treated at hospitals participating in clinical trials (i.e., trial hospitals) were compared with all patients treated at hospitals that did not participate in any clinical trials conducted by AGO-OVAR and NOGGO (i.e., non-trial hospitals) in the specified time period. The primary outcome measure was overall survival time, which was defined as the time interval between the date of diagnosis and the date of death from any cause or the date of the last contact (censored). Survival status was followed up yearly for a period of at least three years after diagnosis. The following patient and disease variables known to be important prognostic factors for ovarian cancer were collected at baseline (see Section 1.3.1 above): patient age at diagnosis, FIGO stage, ECOG performance status, tumor grade, histological subtype, and presence of ascites of more than 500 ml. In addition, presence of comorbidity and history of second malignancies were recorded.

The use of data included in this thesis was approved by the AGO-OVAR and the AGO Ovarian Committee. The study protocol was reviewed and approved by the Ethics Committee of the Medical Faculty of the University of Heidelberg, Germany (study number: S-446/2013).

2.2 Methods

The quality assurance program QS-OVAR follows an observational cross-sectional study design with retrospective data collection and prospective follow up. About a year after initial diagnosis, hospitals submit their ovarian cancer patient data. This time frame typically marks the completion of both surgical treatment and primary chemotherapy, with patients transitioning into the follow-up phase. This thesis utilizes data from three distinct cohorts of patients diagnosed with epithelial ovarian cancer in the years 2001, 2004, and 2008. For the first cohort, survival data were gathered for a minimum of three years post diagnosis, while in the subsequent cohorts, this information was collected for at least four years following initial diagnosis. The results are thus presented separately for QS-OVAR 2001, 2004 and 2008. Changes over time with respect to adherence to treatment guidelines and ovarian cancer survival are evaluated in a descriptive manner. Subgroup analyses for early-stage (FIGO I–IIA) and advanced-stage (FIGO IIB–IV) ovarian cancer are shown as appropriate.

Continuous data are summarized with medians and 25th and 75th percentiles (i.e., the interquartile range [IQR]), while categorical data are summarized with counts and percentages. Descriptive statistics for survival data are estimates from Kaplan-Meier curves (KM, Kaplan and Meier 1958). Median follow-up time was estimated using the reverse KM method. The reverse KM was calculated in the same way as the standard KM but with the status indicator reversed so that patients alive are treated as events and deaths are censored. The follow-up time of a patient that died is then interpreted as the follow-up time that potentially would have been observed had that patient not died (Schemper and Smith 1996). The total effect of institutional research activity on patient survival was estimated by Cox proportional hazards regression, and the standard error of the effect estimate was corrected for potential clustering of patients within hospitals (Therneau 2013). Stage of disease (FIGO I–IIA vs. FIGO IIB–IV), patient age at diagnosis (continuous, in 5 years units), performance status (ECOG > 1 vs. 0/1), ascites (> 500 ml vs. ≤ 500 ml), comorbidity (present vs. none), histology (serous vs. other), and grade (G3/4 vs. G1/2) were included as baseline covariates or factors.

One main question of the present thesis was whether the effect of hospital trial participation on patient survival is mediated through better adherence to treatment guidelines regarding chemotherapy selection and surgical outcome. Thus, in the next step, the effects of research activity on the two binary process variables “optimal surgery” and “optimal chemotherapy” were investigated. Odds ratios for both outcomes were obtained from logistic regression models fitted with generalized estimating equations (GEE) to account for clustering of patients within hospitals (Højsgaard et al. 2006). The relationship between adherence to treatment guidelines and survival was assessed by using Cox proportional hazards model with robust variance estimator (Therneau 2013).

It is crucial to adjust for the potential confounders in the statistical models for the mediators and the outcome. There is confounding in the exposure-outcome relationship if another variable is related to the exposure that also has an influence on the outcome. For example, it is plausible to assume that patients at higher age are less likely to seek care at specialized, research-active institutions, and may, in turn, have a lower chance to receive the therapeutic standard treatment. An analysis that does not take age into account would then be subject to unmeasured confounding in the exposure-mediator relationship which would lead to artificially exaggerated mediator effects. Therefore, the above-mentioned prognostic factors were included into the models to adjust for known baseline confounders.

Causal mediation analysis was performed by using the approach proposed by Lange et al. (2012). This approach is based on the Neyman-Rubin counterfactual framework of causality (Splawa-Neyman 1990; Rubin 1974; Rubin 2004). A counterfactual presents a potential outcome that would have happened in the absence of the cause. For example, either an ovarian cancer patient was treated at a trial hospital or not. However, one does not observe what would have happened to a patient had this patient not been treated at a trial hospital, when she was; and one does not observe what would have happened to a patient had she been treated at a trial hospital, when she was not. By its nature a counterfactual is not observed in real data and, therefore, must be estimated. The approach proposed by Lange et al. (2012) allows decomposition of the total effect of a given exposure *A* on the outcome *Y* into a natural direct effect ($A \rightarrow Y$ in Figure 2) and a natural indirect effect through a mediator *M* ($A \rightarrow M \rightarrow Y$ in Figure 2). In case of a time-to-event outcome *Y*, a binary exposure *A*, a binary mediator *M* and a number of baseline confounders *C*, Lange et al. (2012) showed that unbiased estimates for the direct and indirect effect are obtained from a weighted Cox regression of the time-to-event outcome on *A*, *A** and *C* using a duplicated data set. In the first replication *A** takes the original value of the exposure. In the second replication *A** takes the opposite (counterfactual) value of the exposure. The weights are determined by

$$W_c = \mathrm{P}(M \mid A^*, C) \,/\, \mathrm{P}(M \mid A, C),$$

with $\mathrm{P}(\cdot)$ deriving from a logistic regression of the mediator *M* on the exposure and the baseline confounders (Lange et al. 2012, Appx. 4). Assuming non-informative censoring and proportional hazards, the weighted Cox model yields hazard ratios for *A* and *A** that serve as estimates for the natural direct effect and the natural indirect effect, respectively. The product of the two hazard ratios yields the hazard ratio for the total effect. Standard errors and confidence intervals can be determined using, for example, bootstrap methods.

For the case of multiple mediators (e.g., like in the present study, with the two mediators optimal surgery and optimal chemotherapy), the approach of Lange et al. (2012) is mathematically consistent only in the absence of mediator-mediator interactions. Therefore, one cannot simply assume that the two mediator effects are consistently estimated by running two analyses with a single mediator. It is also crucial that the causal pathways are non-intertwined, that is, the mediators must not have causal effects onto each other (see the "Extended sequential ignorability" condition of Lange et al. 2014, Appx. 3). This assumption requires, among others, mutual conditional independence of the mediators, given the exposure and the confounders. In the present study, this assumption was tested using a logistic regression for Mediator 1 on the exposure and the baseline confounders, and then adding Mediator 2 as a further predictor. The odds ratio for Mediator 2 should ideally be close to unity ("non-significant" with a narrow confidence interval, see Lange et al. 2014, Step 3).

As mentioned above, two potential binary mediators (surgery and chemotherapy) of the effect of hospital research activity on survival are explored in the present thesis (Figure 3). Under the assumption of separate, independent, causal pathways through the two mediators, unbiased point estimates for the natural direct effect and the natural indirect effects related to the two mediators can be obtained by a weighted Cox regression of the outcome on the exposure, the baseline confounders and two additional counterfactual variables A_1* and A_2* that were systematically manipulated in four replicates of the original data (based on Lange et al. 2014, Step 4).

The main analyses were supported by various sensitivity analyses. These sensitivity analyses explored the influence of two other institutional characteristics (hospital volume and care level) and also examined more recent definitions of adherence to treatment guidelines. Further sensitivity analyses were carried out explicitly to examine assumptions underlying the causal mediation analysis. The multiple pathways framework according to Lange et al. (2014) assumes that the two mediators are fulfilled independently of each other, as well as that the two mediators operate separately of each other. Because it is generally difficult to test such assumptions within the same data set, sensitivity analyses were performed to assess the robustness of the results. In the first sensitivity analysis, a single binary mediator was used that reflected optimal adherence to treatment guidelines in the sense that both the chemotherapy and surgery were optimal. In a second sensitivity analysis, the number of fulfilled criteria was counted, that is, the mediator was again a single variable indicating whether none (neither chemotherapy nor surgery), one (optimal chemotherapy or optimal surgery), or both criteria for treatment adherence (optimal chemotherapy and optimal surgery) were met. Finally, in a last sensitivity analysis, possible bias due to non-linear relationships and interactions between exposure, baseline variables and mediators, as well as possible misclassification of the mediators were investigated.

All analyses were conducted in R version 3.1.2 (R Core Team 2013) and were performed according to the intention-to-treat principle, meaning that data from all hospitals and their eligible patients are included. Confidence intervals for mediation effects that account for clustering of patients within hospitals were obtained using simple random cluster sampling and 10,000 bootstrap simulations. In all analyses, results are considered statistically significant if the 95% confidence interval (CI) for the hazard ratio (HR) or odds ratio (OR) do not include 1 or two-sided *P* values are less than 0.05. Owing to the exploratory character of QS-OVAR and because *P* values are reported without adjustment, all reported statistics are to be interpreted in a strictly descriptive way and not in a confirmatory manner.

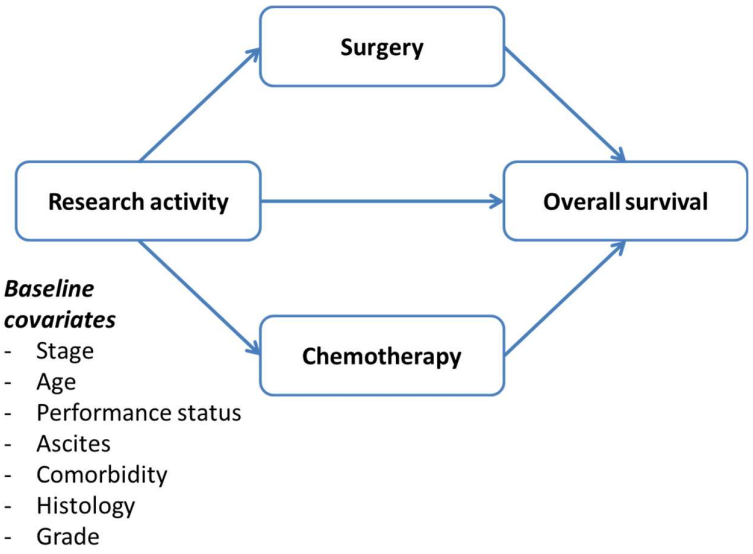

Figure 3: Path diagram relating hospital research activity to two mediators and overall survival time in the presence of known baseline confounders

3 Results

In Phase 1 of each survey, all gynecological departments in Germany were asked to report the number of newly diagnosed patients with epithelial ovarian cancer in the respective year, and about their membership in the two German study groups, AGO-OVAR and NOGGO. In the first cohort, 481 (43%) of the 1123 contacted hospitals provided the required information. Almost the same response rate was obtained for the second cohort (469 of 1109 hospitals, 42%). Finally, more than half of the hospitals that were invited to participate in QS-OVAR 2008 responded to the survey (507 of 974 hospitals, 52%). The response rate in Phase 1 was positively correlated with hospital care level: The higher the care level of the hospital was, the higher was the likelihood of response and participation in Phase 2 (Figure 4). For example, in the last QS-OVAR 2008 cohort, all of the 36 contacted university hospitals responded to the survey. For comparison, there were 61 responders from hospitals with centralized services (91% of 67) and 113 responders from hospitals with specialized care (73% of 155). In contrast, only 41% of the 716 primary care hospitals replied to the survey in 2008. Similarly, almost all of the university hospitals but only about one third of the primary care hospitals responded to the survey in the first two cohorts. Among responders there were up to 20% of hospitals in the primary care sector that had not treated any ovarian cancer patients in the requested time period.

In Phase 2, the responding hospitals were asked to document all patients diagnosed with ovarian cancer in the third quarter of the respective year. This phase was initiated approximately one year after diagnosis. Data from altogether 2120 patients were obtained in Phase 2. There was a remarkable increase in the number of included patients and institutions over the years: QS-OVAR 2001 consisted of 476 patients from 165 hospitals, QS-OVAR 2004 included 763 patients from 245 hospitals, and QS-OVAR 2008 contained data on 881 patients from 240 hospitals.

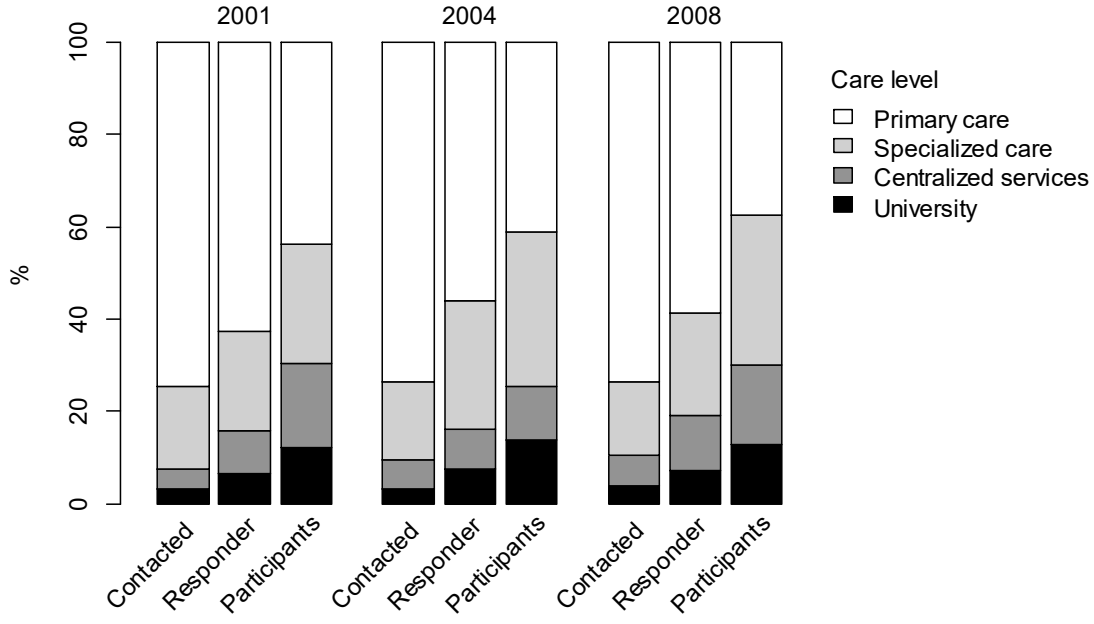

Figure 4: Hospital care level in Phase 1 and Phase 2 of QS-OVAR by year of diagnosis

3.1 Hospital characteristics

In each cohort of QS-OVAR, about 50% of the participating hospitals in Phase 2 were research-active ("trial hospitals", i.e., they participated in AGO or NOGGO trials before or during the third quarter of the respective year). However, only 21% of patients treated in trial hospitals in 2001, 6% of patients in trial hospitals in 2004, and 16% in trial hospitals in 2008 were actually enrolled in prospective randomized trials of the two study groups. The median number of annual ovarian cancer cases was 12 (IQR 8 to 20, range 1 to 90). There was a slight increase from 2001 to 2008 in the number of hospitals with at least 12 patients per year. About 30% of the hospitals were university hospitals or hospitals with centralized services. Compared to Phase 1, a trend towards higher hospital care level was observed in Phase 2 (see responder in Phase 1 vs. participants in Phase 2 in Figure 4). When looking at the participation of the individual hospitals in QS-OVAR over time, there was some overlap between the cohorts: A total of 67 hospitals entered all three cohorts while 123 hospitals participated twice in the survey. In contrast, 203 hospitals participated only once in QS-OVAR. With regard to the latter, 24% of the 203 hospitals were included only in the first cohort, 36% were included only in the second cohort and 40% entered only the third cohort. Further information on hospital characteristics by year of diagnosis is displayed in Table 2.

Table 2: Hospital characteristics

	QS-OVAR 2001 (*N* = 165)		QS-OVAR 2004 (*N* = 245)		QS-OVAR 2008 (*N* = 240)	
	N	(%)	*N*	(%)	*N*	(%)
Research activity						
Trial hospital	80	(48.5)	121	(49.4)	114	(47.5)
Non-trial hospital	85	(51.5)	124	(50.6)	126	(52.5)
Volume						
1–11 patients/year	81	(49.1)	106	(43.3)	97	(40.4)
12+ patients/year	84	(50.9)	139	(56.7)	143	(59.6)
Care level						
University hospital	20	(12.1)	34	(13.9)	31	(12.9)
Hospital with centralized services	30	(18.2)	28	(11.4)	41	(17.1)
Hospital providing specialized care	43	(26.1)	82	(33.5)	78	(32.5)
Primary care hospital	72	(43.6)	101	(41.2)	90	(37.5)

Note: The numbers refer only to hospitals participating in Phase 2 of the respective QS-OVAR.

Bivariate analyses of hospital characteristics indicated that research activity was associated with both higher hospital volume (Figure 5) and higher hospital care level (Figure 6). In particular, in the last cohort, more than 80% of trial hospitals but only 40% of non-trial hospitals treated at least 12 patients with ovarian cancer per year (Figure 5). In all three cohorts, about one fourth of the trial hospitals but only 1 to 2% of non-trial hospitals had a university affiliation. In contrast, more than half of the non-trial hospitals and less than 25% of the trial hospitals were primary care hospitals (Figure 6).

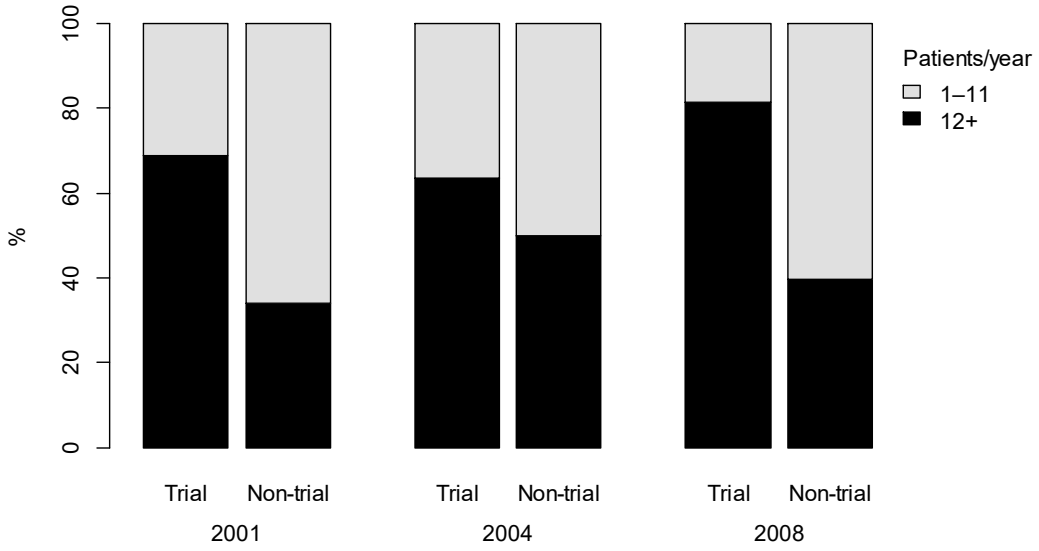

Figure 5: Hospital research activity and hospital volume by year of diagnosis

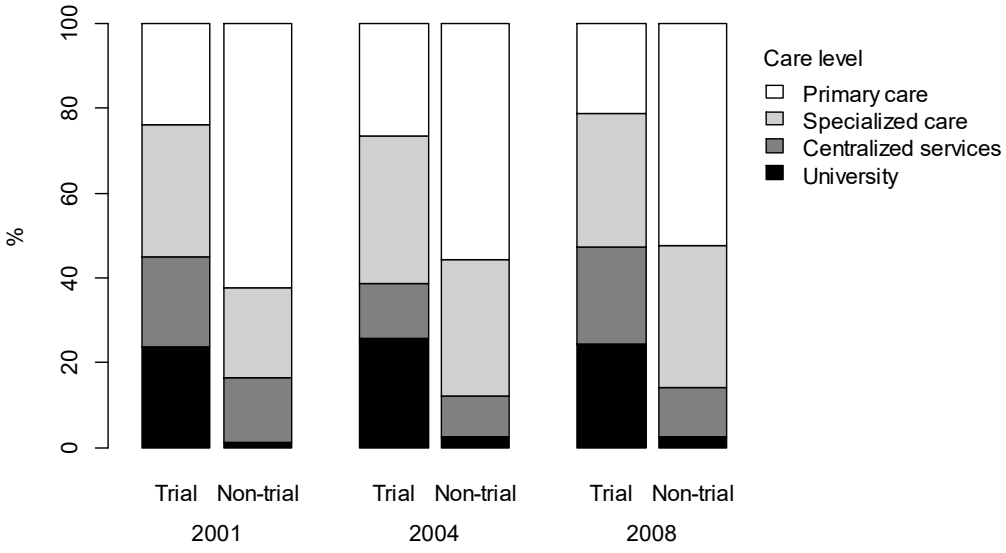

Figure 6: Hospital research activity and hospital care level by year of diagnosis

3.2 Patient and disease characteristics

The median age at diagnosis of ovarian cancer for the entire study population of QS-OVAR was 65 years (IQR 54 to 73), and 17% of the 2120 patients were diagnosed before the age of 50. The majority of the patients (75%) were in good performance status (ECOG 0 or 1). Twenty-one percent of patients were diagnosed with FIGO stage I ovarian cancer, 10% with FIGO stage II, 55% with FIGO stage III, and 15% with FIGO stage IV. The advanced stages of the disease (FIGO IIB–IV) accounted for 77% of patients. About 50% of the tumors were at least moderately differentiated. Serous tumors were the most common histologic subtype (67%), followed by endometrioid (10%) and mucinous (7%) tumors. In contrast, clear-cell tumors were relatively rare (2%). Ascites of more than 500 ml occurred in 46% of patients. Almost 30% of patients suffered from at least one relevant comorbid condition (e.g., cardiovascular disorders, diabetes mellitus, gastrointestinal disorders, pulmonary disease, or chronic kidney disease). Secondary malignancies were reported in 14% of patients. Among these, the most common were breast cancer (41%), endometrial cancer (20%), and colon cancer (11%).

In general, the cohorts were quite similar with regard to patient and disease characteristics. Nevertheless, the following differences were noted: The patients appeared to get older with each cohort. The percentages of elderly patients (age 65 and older) were 46% in 2001, 49% in 2004, and 53% in 2008. In addition, the patients seemed to be diagnosed in more advanced stage of ovarian cancer and with more high-grade tumors from cohort to cohort. In QS-OVAR 2001, about 74% of patients had FIGO stage IIB or higher; the corresponding proportions in QS-OVAR 2004 and 2008 were 76% and 80%, respectively. Similarly, there was a time trend towards higher-grade tumors: Grade 3 or 4 tumors were observed in 45% of patients diagnosed in 2001 and in 46% in 2004, and in almost 57% of patients diagnosed in 2008. Finally, the patients seemed to more likely present with relevant ascites and with comorbidity at diagnosis over the years. While in 2001 about 40% of patients had ascites of more than 500 ml, the two later cohorts included 46% and 48% of patients with relevant ascites at the time of diagnosis. In QS-OVAR 2001, 24% of the patients had at least one comorbid condition, the corresponding proportions of patients in 2004 and 2008 were 27% and 33%, respectively.

Overall, 1258 of the 2120 patients (59%) included in the QS-OVAR program were treated in hospitals participating in clinical trials. In the first cohort of QS-OVAR, 275 of 476 patients were treated in trial hospitals (58%), in the second cohort these were 413 out of 763 patients (54%), and finally, in the third cohort 570 out of 881 patients (65%) were treated in hospitals participating in clinical trials.

Table 3 summarizes the patient and disease characteristics by QS-OVAR cohort and hospital research activity. Patients treated in trial hospitals were generally similar to patients treated in non-trial hospitals. However, some imbalances were observed. For example, in the QS-OVAR 2001 cohort, patients treated in trial hospitals had higher FIGO stages and poorer differentiated tumors than patients treated in non-trial hospitals (i.e., 80% of patients in trial hospitals were diagnosed with FIGO stage IIB or higher ovarian cancer compared to 66% of patients in non-trial hospitals; 50% of patients in trial hospitals had G3 or G4 tumors compared to 39% of patients in non-trial hospitals). Similarly, the first two cohorts included more patients with serous histology in trial hospitals than in non-trial hospitals. In contrast, non-trial hospitals treated more patients with less favorable prognostic factors in QS-OVAR 2008. In this latest cohort, there were more elderly patients and slightly more patients with worse ECOG performance status in non-trial hospitals than in trial hospitals. Similarly, in 2008 the proportion of patients with ascites of more than 500 ml was higher in non-trial hospitals when compared to hospitals participating in trials (for details, see Table 3).

Table 3: Patient and disease characteristics

	QS-OVAR 2001 hospitals						QS-OVAR 2004 hospitals						QS-OVAR 2008 hospitals					
	Trial (*N* = 275)		Non-trial (*N* = 201)		All (*N* = 476)		Trial (*N* = 413)		Non-trial (*N* = 350)		All (*N* = 763)		Trial (*N* = 570)		Non-trial (*N* = 311)		All (*N* = 881)	
Age in years																		
Median (IQR)	63	(54–73)	64	(53–73)	64	(54–73)	65	(55–73)	65	(55–74)	65	(55–73)	65	(53–73)	68	(55–74)	66	(54–73)
FIGO stage																		
IA, *N* (%)	22	(8.0)	28	(13.9)	50	(10.5)	37	(9.0)	36	(10.3)	73	(9.6)	56	(9.8)	26	(8.4)	82	(9.3)
IB	4	(1.5)	5	(2.5)	9	(1.9)	2	(0.5)	6	(1.7)	8	(1.0)	5	(0.9)	3	(1.0)	8	(0.9)
IC	24	(8.7)	32	(15.9)	56	(11.8)	45	(10.9)	35	(10.0)	80	(10.5)	48	(8.4)	27	(8.7)	75	(8.5)
IIA	6	(2.2)	3	(1.5)	9	(1.9)	11	(2.7)	11	(3.1)	22	(2.9)	12	(2.1)	4	(1.3)	16	(1.8)
IIB	6	(2.2)	4	(2.0)	10	(2.1)	12	(2.9)	11	(3.1)	23	(3.0)	20	(3.5)	15	(4.8)	35	(4.0)
IIC	11	(4.0)	10	(5.0)	21	(4.4)	9	(2.2)	19	(5.4)	28	(3.7)	20	(3.5)	17	(5.5)	37	(4.2)
IIIA	10	(3.6)	3	(1.5)	13	(2.7)	8	(1.9)	2	(0.6)	10	(1.3)	12	(2.1)	3	(1.0)	15	(1.7)
IIIB	12	(4.4)	23	(11.4)	35	(7.4)	23	(5.6)	26	(7.4)	49	(6.4)	30	(5.3)	18	(5.8)	48	(5.4)
IIIC	144	(52.4)	75	(37.3)	219	(46.0)	194	(47.0)	139	(39.7)	333	(43.6)	286	(50.2)	156	(50.2)	442	(50.2)
IV	36	(13.1)	18	(9.0)	54	(11.3)	72	(17.4)	65	(18.6)	137	(18.0)	81	(14.2)	42	(13.5)	123	(14.0)
ECOG performance status																		
0, *N* (%)	110	(40.0)	91	(45.3)	201	(42.2)	163	(39.5)	133	(38.0)	296	(38.8)	229	(40.2)	101	(32.5)	330	(37.5)
1	106	(38.5)	67	(33.3)	173	(36.3)	147	(35.6)	121	(34.6)	268	(35.1)	217	(38.1)	110	(35.4)	327	(37.1)
2	36	(13.1)	32	(15.9)	68	(14.3)	54	(13.1)	52	(14.9)	106	(13.9)	66	(11.6)	47	(15.1)	113	(12.8)
3	22	(8.0)	9	(4.5)	31	(6.5)	39	(9.4)	36	(10.3)	75	(9.8)	37	(6.5)	33	(10.6)	70	(7.9)
4	1	(0.4)	2	(1.0)	3	(0.6)	2	(0.5)	3	(0.9)	5	(0.7)	4	(0.7)	0	(0.0)	4	(0.5)
NA	0	(0.0)	0	(0.0)	0.0	(0.0)	8	(1.9)	5	(1.4)	13	(1.7)	17	(3.0)	20	(6.4)	37	(4.2)

Note: The dashed line indicates the boundary at which the variable was dichotomized in the prognostic models.

Table 3: Patient and disease characteristics

	QS-OVAR 2001 hospitals						QS-OVAR 2004 hospitals						QS-OVAR 2008 hospitals					
	Trial (*N* = 275)		Non-trial (*N* = 201)		All (*N* = 476)		Trial (*N* = 413)		Non-trial (*N* = 350)		All (*N* = 763)		Trial (*N* = 570)		Non-trial (*N* = 311)		All (*N* = 881)	
Tumor grade																		
G4, *N* (%)	1	(0.4)	0	(0.0)	1	(0.2)	2	(0.5)	4	(1.1)	6	(0.8)	3	(0.5)	4	(1.3)	7	(0.8)
G3	136	(49.5)	78	(38.8)	214	(45.0)	190	(46.0)	158	(45.1)	348	(45.6)	321	(56.3)	171	(55.0)	492	(55.8)
G2	98	(35.6)	68	(33.8)	166	(34.9)	161	(39.0)	126	(36.0)	287	(37.6)	172	(30.2)	98	(31.5)	270	(30.6)
G1	33	(12.0)	31	(15.4)	64	(13.4)	26	(6.3)	29	(8.3)	55	(7.2)	38	(6.7)	17	(5.5)	55	(6.2)
GX	7	(2.5)	24	(11.9)	31	(6.5)	34	(8.2)	33	(9.4)	67	(8.8)	36	(6.3)	21	(6.8)	57	(6.5)
Histology																		
Serous, *N* (%)	196	(71.3)	133	(66.2)	329	(69.1)	277	(67.1)	204	(58.3)	481	(63.0)	397	(69.6)	219	(70.4)	616	(69.9)
Endometrioid	30	(10.9)	20	(10.0)	50	(10.5)	39	(9.4)	43	(12.3)	82	(10.7)	52	(9.1)	31	(10.0)	83	(9.4)
Mucinous	20	(7.3)	20	(10.0)	40	(8.4)	26	(6.3)	30	(8.6)	56	(7.3)	32	(5.6)	24	(7.7)	56	(6.4)
Other	29	(10.5)	28	(13.9)	57	(12.0)	71	(17.2)	73	(20.9)	144	(18.9)	89	(15.6)	37	(11.9)	126	(14.3)
Ascites																		
> 500 ml, *N* (%)	111	(40.4)	81	(40.3)	192	(40.3)	191	(46.2)	159	(45.4)	350	(45.9)	259	(45.4)	167	(53.7)	426	(48.4)
≤ 500 ml	163	(59.3)	118	(58.7)	281	(59.0)	218	(52.8)	188	(53.7)	406	(53.2)	305	(53.5)	143	(46.0)	448	(50.9)
NA	1	(0.4)	2	(1.0)	3	(0.6)	4	(1.0)	3	(0.9)	7	(0.9)	6	(1.1)	1	(0.3)	7	(0.8)
Comorbidity																		
Present, *N* (%)	65	(23.6)	51	(25.4)	116	(24.4)	117	(28.3)	85	(24.3)	202	(26.5)	195	(34.2)	99	(31.8)	294	(33.4)
None	210	(76.4)	150	(74.6)	360	(75.6)	296	(71.7)	265	(75.7)	561	(73.5)	375	(65.8)	212	(68.2)	587	(66.6)
Second malignancy																		
Yes, *N* (%)	34	(12.4)	34	(16.9)	68	(14.3)	62	(15.0)	44	(12.6)	106	(13.9)	74	(13.0)	38	(12.2)	112	(12.7)
No	241	(87.6)	167	(83.1)	408	(85.7)	351	(85.0)	306	(87.4)	657	(86.1)	496	(87.0)	273	(87.8)	769	(87.3)

Note: The dashed line indicates the boundary at which the variable was dichotomized in the prognostic models.

3.3 Adherence to treatment guidelines

Adherence to treatment guidelines was defined as stage-appropriate surgery and chemotherapy (see Section 2.1). Therefore, treatment data will be presented separately for patients with early-stage and advanced-stage disease. Altogether, 488 patients were diagnosed with early (FIGO I–IIA) ovarian cancer: 124 patients in the first cohort, 183 patients in the second cohort, and 181 patients in the third cohort of QS-OVAR. In contrast, 1632 patients had advanced (FIGO IIB–IV) ovarian cancer at the time of diagnosis: 352 patients in QS-OVAR 2001, 580 patients in 2004, and 700 patients in 2008.

3.3.1 Adherence to surgical guidelines

3.3.1.1 Early-stage ovarian cancer

The surgical treatment standard for early ovarian cancer was defined as accurate staging including vertical laparotomy, total abdominal hysterectomy, bilateral salpingo-oophorectomy with removal of all tumor tissue, omentectomy, peritoneal sampling, cytology, as well as pelvic and para-aortic lymph node staging (see Section 2.1 above). Please note, however, that total abdominal hysterectomy and bilateral salpingo-oophorectomy were not considered mandatory in patients with highly differentiated FIGO IA tumors and an option for fertility sparing surgery. Among the 488 patients diagnosed with FIGO I–IIA ovarian cancer, 20 patients (4%) were younger than 50 years and had FIGO IA G1 tumors: seven patients in 2001, 11 patients in 2004, and two patients in 2008.

Overall, only 102 of 488 patients (21%) with early-stage ovarian cancer received surgical staging that would be considered "complete". One staging step was missing in 108 (22%) patients, two staging procedures were omitted in 96 (20%) patients, three staging steps were missing in 61 (13%) patients, four in 50 (10%), five in 35 (7%), six in 21 (4%), seven in 10 (2%), and eight procedures were lacking in four patients (1%). There was one patient without any staging of disease. When looking into single staging procedures, some of them were more frequently missing than others. For example, surgical lymph node staging was not done at all in 37% of the patients. Pelvic lymph node dissection was missing in 38% of patients, and para-aortal lymph node dissection was missing in even 57% of patients. Two of the other most frequently omitted staging steps were peritoneal biopsies and cytology that were not performed in 56% and 28% of patients, respectively. It is of note that both procedures neither need special surgical skills nor are they associated with remarkable burden for the patient.

In QS-OVAR 2001, all of the required staging steps were done in only 6% of the 124 patients with early ovarian cancer. This proportion increased to 19% in 2004 and to almost 34% in 2008. A similar positive time trend was observed with regard to the weaker definition of adherence to staging guidelines for FIGO I–IIA that allowed omission of one staging step. According to this definition, staging with maximally one procedure missing was conducted in only 23% of patients in 2001. This proportion increased to 39% in 2004 and to even 61% in 2008. When looking into single staging procedures by cohort, lymph node staging was not done at all in almost half of the patients (47%) in 2001. This percentage decreased to 37% in 2004 and to 30% in 2008. In 2001, pelvic and para-aortic lymph node dissections were missing in 49% and in 77% of patients, respectively. These percentages decreased to 37% and 63% in QS-OVAR 2004, and to 32% and 37% in QS-OVAR 2008, respectively. A similar positive time trend towards higher completion rate was observed for peritoneal biopsies and cytology (for more details see Table 4A).

Patients in trial hospitals had a slightly higher chance of receiving more complete staging than those treated in hospitals not participating in trials (Table 4A, Figure 7). In QS-OVAR 2001, 5 of 56 patients (9%) received complete surgical staging in trial hospitals and 2 of 68 patients (3%) in non-trial hospitals. In 2004, these were 21 of 95 (22%) patients in trial hospitals compared to 13 of 88 patients (15%) in non-trial hospitals. Finally, in 2008, 44 of 121 patients (36%) in trial hospitals and 17 of 60 patients (28%) in non-trial hospitals were completely staged. When surgical staging with none or only one missing staging step was considered, positive effects of institutional participation in clinical trials were only observed in the first and the last cohort. In 2001, staging with maximally one omitted procedure was achieved in 32% of patients in trial hospitals compared to 15% in non-trial hospitals. The corresponding percentages in 2008 were 65% in trial hospitals and 53% in non-trial hospitals, respectively. In contrast, in QS-OVAR 2004, the percentage of patients with maximally one staging procedure missing was even slightly lower in trial hospitals (36%) compared to non-trial hospitals (42%).

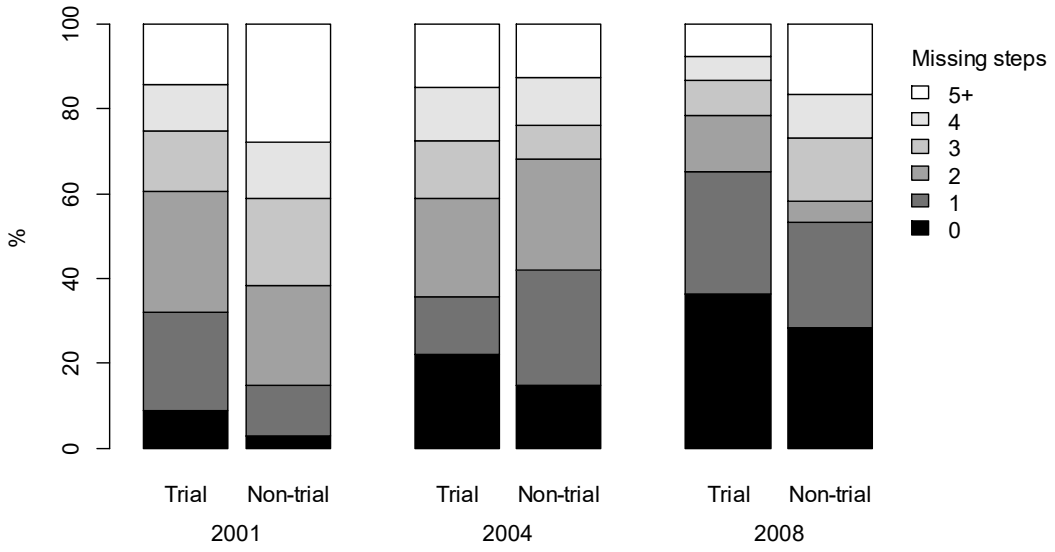

Figure 7: Missing staging steps in early-stage ovarian cancer in trial and non-trial hospitals by year of diagnosis

Figures 8 and 9 give an overview of the number of omitted staging steps according to hospital volume and hospital care level. As can be seen from Figure 8, almost the same staging pattern as for hospital research activity was observed for hospital volume. Patients in high-volume hospitals treating at least 12 patients per year had a higher chance of receiving more complete staging than patients treated in low-volume hospitals. Similar effects were found when maximally one of the nine staging steps was defined as acceptable to be omitted. As can be seen from Figure 9, the staging pattern was less clear when hospital care level was evaluated. Even though, both complete staging and staging with maximally one procedure missing were done more frequently in university hospitals in all three QS-OVAR cohorts, patients in hospitals providing specialized care also seemed to have a quite high chance of receiving accurate staging, at least in 2004 and 2008.

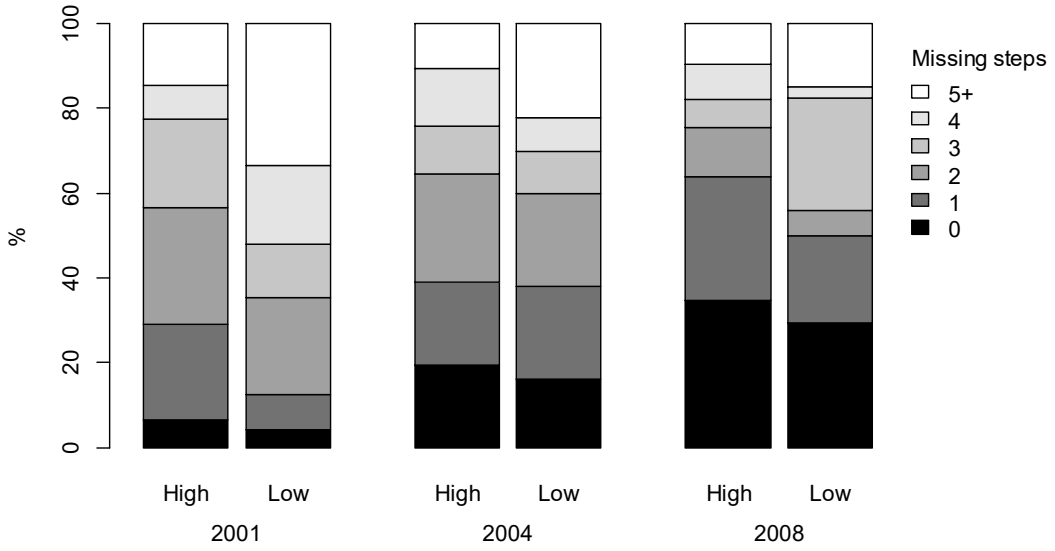

Figure 8: Missing staging steps in early-stage ovarian cancer in high-volume (12+ patients/year) and low-volume (1–11 patients/year) hospitals by year of diagnosis

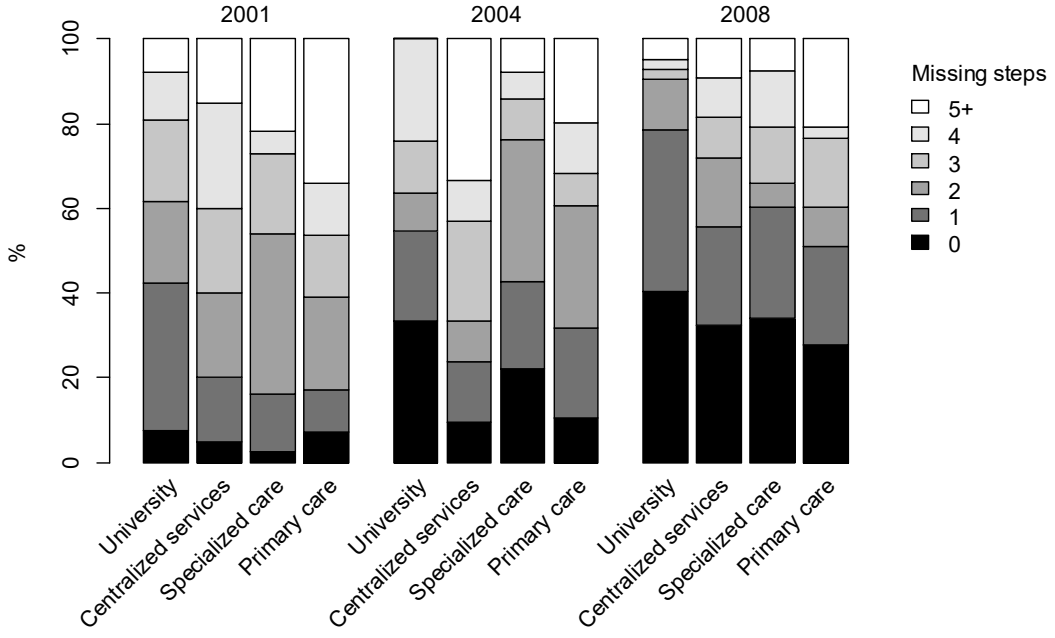

Figure 9: Missing staging steps in early-stage ovarian cancer in university hospitals, hospitals with centralized services and maximum care, hospitals providing specialized care and primary care hospitals by year of diagnosis

3.3.1.2 Advanced-stage ovarian cancer

Primary cytoreductive surgery is the treatment of choice for ovarian cancer. Because the result of this surgery is an important and at the same time a modifiable prognostic factor, preferably no macroscopic residual disease should be the surgical goal to achieve the best prognosis. In this thesis, tumor residuals up to 10 mm largest diameter were considered as "optimal" surgical outcome according to former treatment guidelines. In addition, "complete debulking" was considered in sensitivity analysis to take into account the more recent guidelines that define the primary objective of debulking surgery in FIGO IIB–IV as the complete removal of all visible disease (see Section 2.1).

Complete debulking with no tumor residuals was achieved in altogether 38% of the 1632 patients with FIGO stage IIB–IV (33% in 2001, 37% in 2004, and 42% in 2008). Tumor residuals between 1 and 10 mm were left in 24% of patients (28% in 2001, 23% in 2004, and 23% in 2008). Postoperative tumor residuals larger than 10 mm were left in 37% of patients (39% in 2001, 40% in 2004, and 35% in 2008). Four patients in QS-OVAR 2004 were not treated surgically. In further analysis, these patients were handled as if they had tumor residuals of more than 10 mm (for more details see Table 4B).

Figure 10 shows the surgical outcome in patients with advanced ovarian cancer treated in trial hospitals and non-trial hospitals. Debulking was more often "optimal" (i.e., no visible tumor or residuals between 1 and 10 mm) in research-active hospitals. In QS-OVAR 2001, 46% of patients in non-trial hospitals had tumor residuals larger than 10 mm in contrast to 34% of patients in trial hospitals, in QS-OVAR 2004 these were 47% versus 35%, and in QS-OVAR 2008 the proportions were 43% in non-trial hospitals and 30% in trial hospitals (see also Table 4B). Almost the same pattern was observed for hospital volume: Patients in high-volume hospitals had a higher chance of receiving more complete debulking than patients treated in low-volume hospitals (Figure 11). The pattern was again less consistent for different hospital care levels (Figure 12).

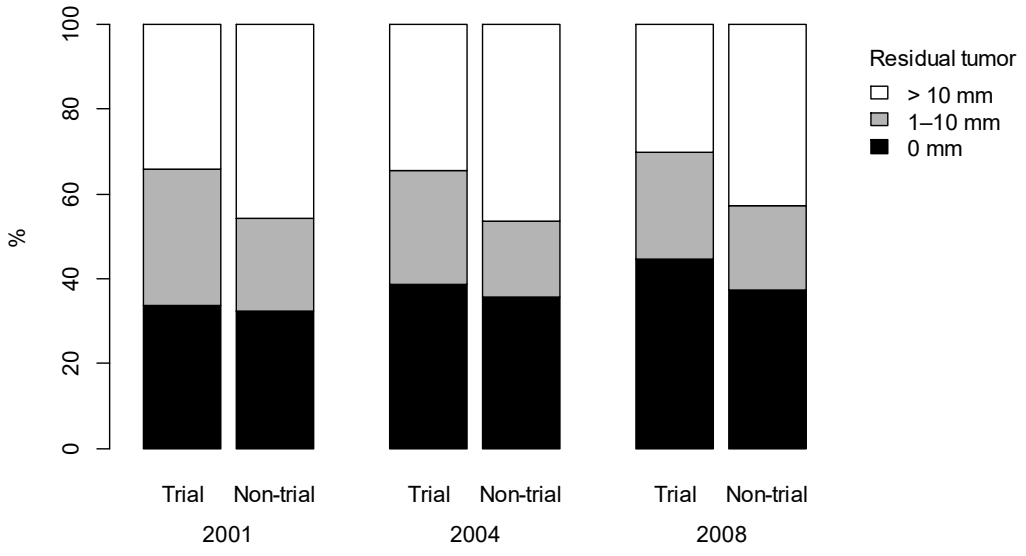

Figure 10: Surgical outcome in advanced-stage ovarian cancer in trial and non-trial hospitals by year of diagnosis

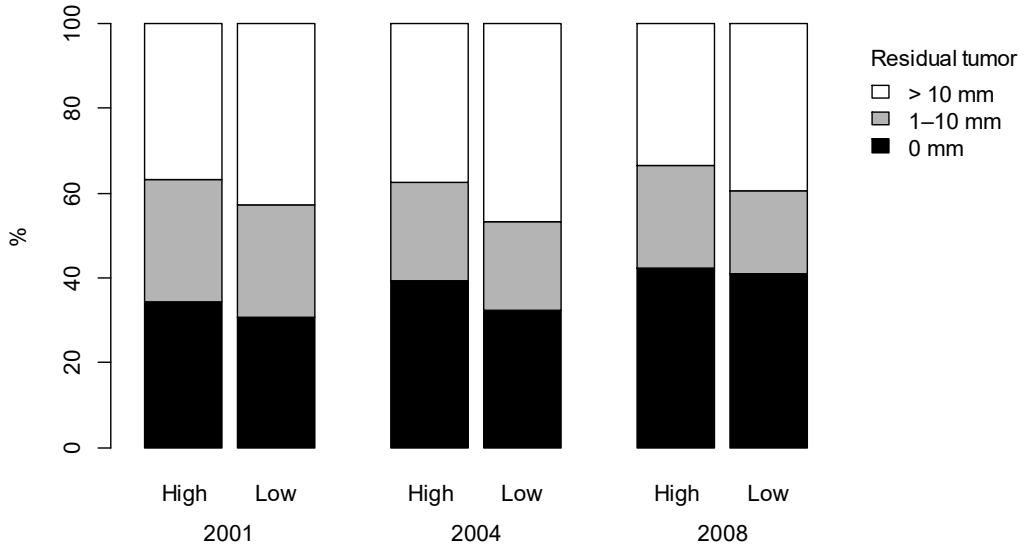

Figure 11: Surgical outcome in advanced-stage ovarian cancer in high-volume (12+ patients/year) and low-volume (1–11 patients/year) hospitals by year of diagnosis

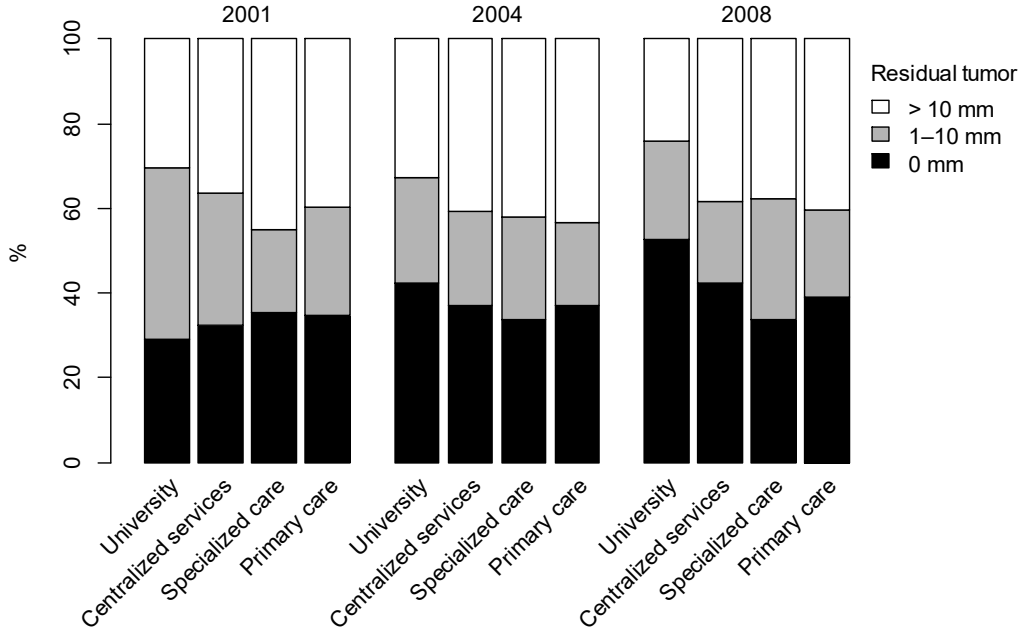

Figure 12: Surgical outcome in advanced-stage ovarian cancer in university hospitals, hospitals with centralized services and maximum care, hospitals providing specialized care and primary care hospitals by year of diagnosis

3.3.1.3 Model for adherence to surgical guidelines

In the next step, the effect of hospital research activity on the appropriateness of surgery was investigated in more detail. Optimal staging in FIGO I–IIA and optimal debulking in FIGO IIB–IV were used as outcome for the primary analysis. In a sensitivity analysis, the stronger "complete" definition of surgery with all nine staging steps done in early-stage disease and with complete removal of all visible tumor in advanced-stage ovarian cancer was considered (see Section 3.5). Logistic regression models were fitted with generalized estimating equations (GEE) to account for clustering of patients within hospitals. Patients of all stages of disease were pooled within cohorts and the predictors of adherence to surgical treatment guidelines were determined for each of the three QS-OVAR cohorts. In addition to trial participation at hospital level, baseline covariates at patient level were stage of disease (FIGO I–IIA vs. FIGO IIB–IV), patient age at diagnosis (continuous, in units of 5 years), ECOG performance status (> 1 vs. 0/1), clinically relevant amount of ascites (> 500 ml vs. ≤ 500 ml), comorbidity (present vs. none), tumor histology (serous vs. other), and tumor grade (G3/4 vs. G1/2).

Table 5 shows the predictors of optimal surgical outcome. It presents the total number of patients for each predictor category, the number of patients for whom an optimal surgical outcome according to guidelines was achieved, and odds ratios with their 95% confidence intervals from logistic regression models. In all three QS-OVAR cohorts, the patients who were treated in trial hospitals were significantly more likely to receive optimal surgical treatment than those patients who were treated in non-trial hospitals. The adjusted odds ratios were 1.83 (95% CI 1.14 to 2.93) in QS-OVAR 2001, and 1.44 in QS-OVAR 2004 (95% CI 1.02 to 2.03) as well as in QS-OVAR 2008 (95% CI 1.00 to 2.07). Patients with advanced-stage disease (FIGO IIB or higher) were more likely to receive optimal surgical treatment than patients with early stage of disease (FIGO IIA or lower). Patient age at diagnosis, performance status, and presence of ascites turned out to be important independent predictors of adherence to surgical guidelines. Compared to younger patients, older patients were significantly less likely to be optimally treated surgically. Similarly, worse performance status and clinically relevant amount of ascites were negatively associated with optimal surgical treatment in all three cohorts of QS-OVAR. Finally, a negative relationship between surgical outcome and predictor was observed for women with comorbid conditions.

Table 4A: Staging in FIGO I–IIA ovarian cancer

Staging	QS-OVAR 2001 hospitals			QS-OVAR 2004 hospitals			QS-OVAR 2008 hospitals		
	Trial (*N* = 56)	Non-trial (*N* = 68)	All (*N* = 124)	Trial (*N* = 95)	Non-trial (*N* = 88)	All (*N* = 183)	Trial (*N* = 121)	Non-trial (*N* = 60)	All (*N* = 181)
Missing steps									
Vertical laparotomy, *N* (%)	9 (16.1)	13 (19.1)	22 (17.7)	9 (9.5)	12 (13.6)	21 (11.5)	5 (4.1)	3 (5.0)	8 (4.4)
Total abdominal hysterectomy*	7 (12.5)	8 (11.8)	15 (12.1)	8 (8.4)	4 (4.5)	12 (6.6)	8 (6.6)	9 (15.0)	17 (9.4)
Bilateral salpingo-oophorectomy*	4 (7.1)	6 (8.8)	10 (8.1)	3 (3.2)	3 (3.4)	6 (3.3)	2 (1.7)	2 (3.3)	4 (2.2)
Removal of all tumor tissue	1 (1.8)	7 (10.3)	8 (6.5)	2 (2.1)	3 (3.4)	5 (2.7)	7 (5.8)	1 (1.7)	8 (4.4)
Omentectomy	15 (26.8)	18 (26.5)	33 (26.6)	18 (18.9)	14 (15.9)	32 (17.5)	13 (10.7)	13 (21.7)	26 (14.4)
Peritoneal sampling	33 (58.9)	52 (76.5)	85 (68.5)	57 (60.0)	49 (55.7)	106 (57.9)	49 (40.5)	31 (51.7)	80 (44.2)
Cytology	13 (23.2)	31 (45.6)	44 (35.5)	32 (33.7)	18 (20.5)	50 (27.3)	27 (22.3)	14 (23.3)	41 (22.7)
Pelvic lymph node dissection	25 (44.6)	35 (51.5)	60 (48.4)	34 (35.8)	33 (37.5)	67 (36.6)	32 (26.4)	25 (41.7)	57 (31.5)
Para-aortic lymph node dissection	37 (66.1)	58 (85.3)	95 (76.6)	60 (63.2)	56 (63.6)	116 (63.4)	38 (31.4)	29 (48.3)	67 (37.0)
Completeness									
0 steps missing ("complete"), *N* (%)	5 (8.9)	2 (2.9)	7 (5.6)	21 (22.1)	13 (14.8)	34 (18.6)	44 (36.4)	17 (28.3)	61 (33.7)
1 step missing	13 (23.2)	8 (11.8)	21 (16.9)	13 (13.7)	24 (27.3)	37 (20.2)	35 (28.9)	15 (25.0)	50 (27.6)
2+ steps missing	38 (67.9)	58 (85.3)	96 (77.4)	61 (64.2)	51 (58.0)	112 (61.2)	42 (34.7)	28 (46.7)	70 (38.7)

Note: *Not mandatory in patients < 50 years with FIGO IA G1 tumors.

Table 4B: Surgical debulking in FIGO IIB–IV ovarian cancer

Debulking (tumor residual)	QS-OVAR 2001 hospitals			QS-OVAR 2004 hospitals			QS-OVAR 2008 hospitals		
	Trial (*N* = 219)	Non-trial (*N* = 133)	All (*N* = 352)	Trial (*N* = 318)	Non-trial (*N* = 262)	All (*N* = 580)	Trial (*N* = 449)	Non-trial (*N* = 251)	All (*N* = 700)
0 mm ("complete"), *N* (%)	74 (33.8)	43 (32.3)	117 (33.2)	122 (38.4)	93 (35.5)	215 (37.1)	201 (44.8)	94 (37.5)	295 (42.1)
1 to 10 mm	70 (32.0)	29 (21.8)	99 (28.1)	85 (26.7)	46 (17.6)	131 (22.6)	113 (25.2)	50 (19.9)	163 (23.3)
> 10 mm	75 (34.2)	61 (45.9)	136 (38.6)	109 (34.3)	121 (46.2)	230 (39.7)	135 (30.1)	107 (42.6)	242 (34.6)
No surgery	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.6)	2 (0.8)	4 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)

Table 5: Predictors of optimal surgery*

	QS-OVAR 2001					QS-OVAR 2004					QS-OVAR 2008				
	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI
Research Activity															
Non-trial hospital	201	82	(40.8)	1	Reference	350	176	(50.3)	1	Reference	311	176	(56.6)	1	Reference
Trial hospital	275	162	(58.9)	1.83	1.14 to 2.93	413	241	(58.4)	1.44	1.02 to 2.03	570	393	(68.9)	1.44	1.00 to 2.07
FIGO stage															
I–IIA	124	28	(22.6)	1	Reference	183	71	(38.8)	1	Reference	181	111	(61.3)	1	Reference
IIB–IV	352	216	(61.4)	11.88	5.81 to 24.27	580	346	(59.7)	5.39	3.33 to 8.74	700	458	(65.4)	3.11	1.94 to 5.00
Age (continuous)															
per 5 years	476	244	(51.3)	0.85	0.78 to 0.93	763	417	(54.7)	0.84	0.78 to 0.90	881	569	(64.6)	0.80	0.75 to 0.87
ECOG performance															
0/1	374	212	(56.7)	1	Reference	564	344	(61.0)	1	Reference	657	479	(72.9)	1	Reference
> 1	102	32	(31.4)	0.40	0.22 to 0.74	199	73	(36.7)	0.51	0.34 to 0.77	224	90	(40.2)	0.45	0.30 to 0.69
Ascites															
≤ 500 ml	284	150	(52.8)	1	Reference	413	245	(59.3)	1	Reference	455	347	(76.3)	1	Reference
> 500 ml	192	94	(49.0)	0.45	0.28 to 0.73	350	172	(49.1)	0.41	0.28 to 0.61	426	222	(52.1)	0.25	0.18 to 0.35
Comorbidity															
None	360	198	(55.0)	1	Reference	561	334	(59.5)	1	Reference	587	414	(70.5)	1	Reference
Present	116	46	(39.7)	0.73	0.42 to 1.27	202	83	(41.1)	0.63	0.43 to 0.93	294	155	(52.7)	0.71	0.49 to 1.01
Histology															
Other	147	71	(48.3)	1	Reference	282	153	(54.3)	1	Reference	265	165	(62.3)	1	Reference
Serous	329	173	(52.6)	0.84	0.51 to 1.39	481	264	(54.9)	0.77	0.54 to 1.08	616	404	(65.6)	1.25	0.86 to 1.83
Grade															
G1/G2	261	125	(47.9)	1	Reference	409	216	(52.8)	1	Reference	382	243	(63.6)	1	Reference
G3/G4	215	119	(55.3)	0.96	0.60 to 1.54	354	201	(56.8)	1.07	0.77 to 1.49	499	326	(65.3)	1.01	0.73 to 1.39

Note: *Optimal surgery: max. 1 missing staging item (FIGO I–IIA)/max. 10 mm tumor residual (FIGO IIB–IV)

3.3.2 Adherence to chemotherapy guidelines

3.3.2.1 Early-stage ovarian cancer

Adjuvant platinum-based chemotherapy was regarded standard care for patients with early ovarian cancer, except for patients with FIGO IA G1 tumors. Altogether, 50 patients had highly differentiated FIGO IA tumors and should, therefore, not have received any chemotherapy. In line with this recommendation, 45 patients were not treated chemotherapeutically (Table 6A). Nevertheless, two patients in 2001 and three patients in 2004 with FIGO IA G1 tumors received some chemotherapy and were thus considered as overtreated against guidelines. Of the remaining 438 patients with early-stage disease who fulfilled the criteria for adjuvant chemotherapy, only 310 (71%) actually received platinum-based treatment (64% of patients in 2001, 68% in 2004, and 78% in 2008, Table 6B). Non-platinum chemotherapy was administered to a total of seven patients. Finally, 121 patients did not receive any adjuvant chemotherapy (35% of patients in 2001, 30% in 2004, and 21% in 2008). Taken together, 355 (73%) of the 488 patients with early-stage disease were treated according to chemotherapy guidelines. The proportion of patients receiving chemotherapy standard increased from 68% and 69% in 2001 respectively 2004 to 80% in QS-OVAR 2008.

Figure 13 shows the adherence to chemotherapy guidelines for early ovarian cancer according to institutional research activity by QS-OVAR cohort. In the first two cohorts, patients in trial hospitals seemed to have a slightly higher chance of being treated according to guidelines than patients in non-trial hospitals (71% vs. 65% in QS-OVAR 2001, 73% vs. 66% in QS-OVAR 2004). However, in QS-OVAR 2008, there was no apparent difference between trial and non-trial hospitals in administration of chemotherapy (see also Table 6A/B).

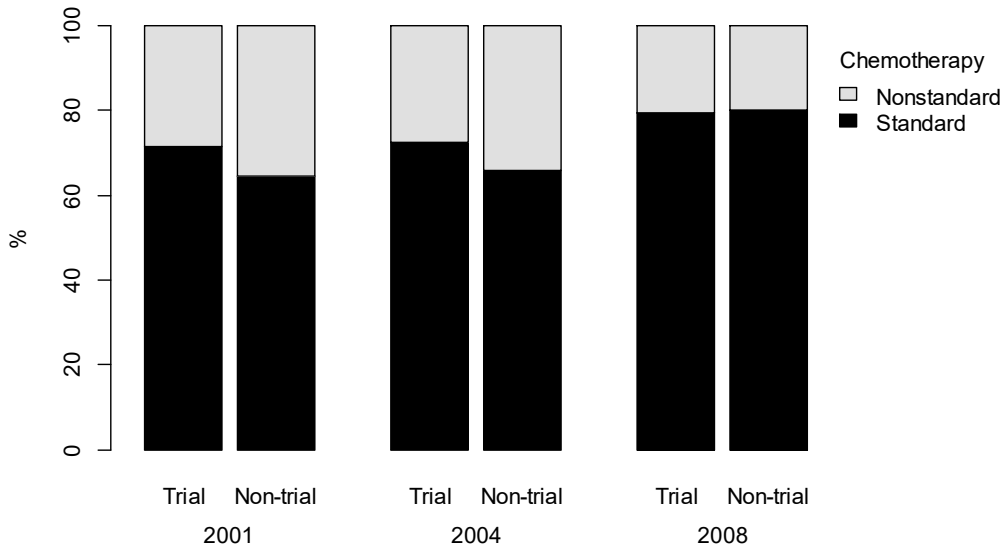

Figure 13: Chemotherapy in early-stage ovarian cancer in trial and non-trial hospitals by year of diagnosis

Patients in high-volume hospitals received chemotherapy standard more often than patients in low-volume hospitals but the effects were only observed in two of the three cohorts (Figure 14). Finally, the pattern was even less consistent when adherence to chemotherapy guidelines was assessed for different hospital care levels (Figure 15).

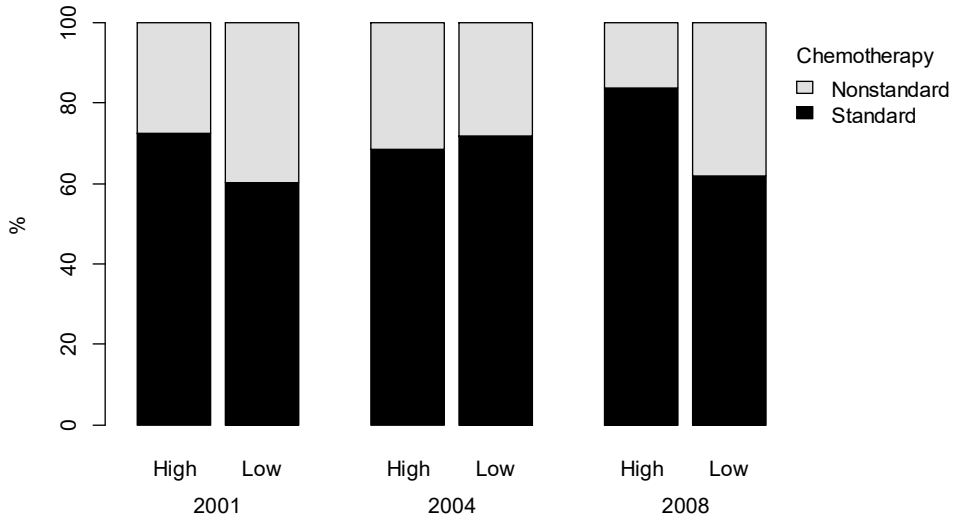

Figure 14: Chemotherapy in early-stage ovarian cancer in high-volume (12+ patients/year) and low-volume (1–11 patients/year) hospitals by year of diagnosis

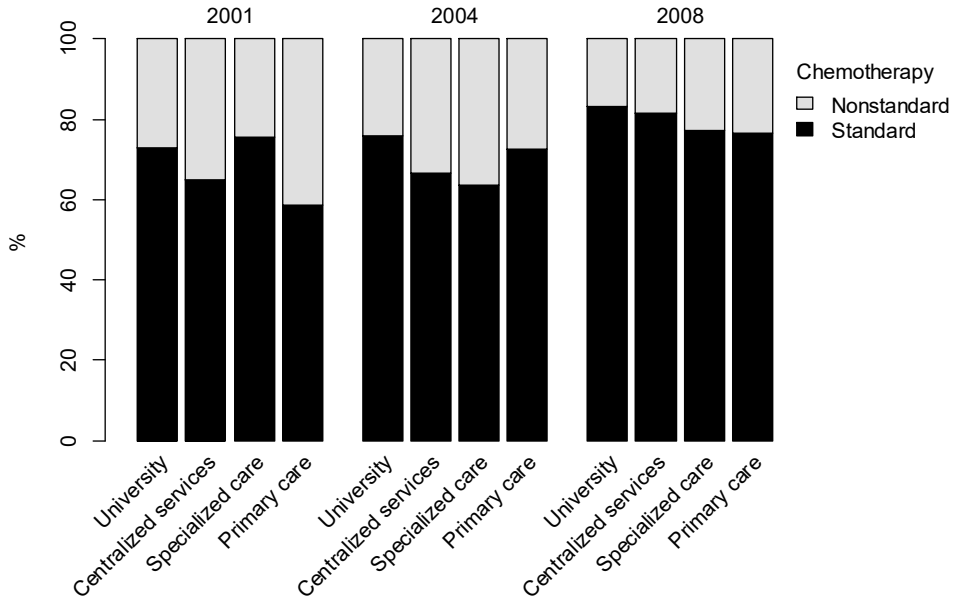

Figure 15: Chemotherapy in early-stage ovarian cancer in university hospitals, hospitals with centralized services and maximum care, hospitals providing specialized care and primary care hospitals by year of diagnosis

3.3.2.2 Advanced-stage ovarian cancer

Traditionally, ovarian cancer has been treated with platinum-based drugs. Cisplatin followed by carboplatin-based combinations with taxanes have been the chemotherapy standard for advanced ovarian cancer for more than 20 years. Accordingly, in this thesis, any combination of platinum and taxane was considered adherent to treatment guidelines in patients with FIGO stage IIB–IV disease.

Of the 1632 patients with advanced-stage disease, 231 (14%) did not receive any chemotherapy. A platinum–taxane combination was administered to 83% of the 1401 patients with chemotherapy. Among the patients with chemotherapy, the proportion of patients receiving platinum-taxane chemotherapy increased slightly from 78% in 2001 to 83% in 2004 and 87% in 2008. Platinum only was given to a total of 13% of patients with chemotherapy (16% in 2001, 13% in 2004, and 11% in 2008), and about 4% of the patients with chemotherapy received other agents as first-line chemotherapy. Altogether, 72% of the 1632 patients with advanced ovarian cancer were treated according to chemotherapy guidelines: 66% in 2001, 70% in 2004, and 76% in 2008.

In all three cohorts, patients in trial hospitals had a higher chance of receiving the therapeutic standard (platinum-taxane chemotherapy) compared to patients in hospitals not participating in clinical studies (70% vs. 59% in QS-OVAR 2001, 75% vs. 65% in 2004, and 77% vs. 73% in 2008, see Table 6C, Figure 16). However, here again, the pattern was less consistent for the other two institutional characteristics, hospital volume and hospital care level (Figures 17 and 18).

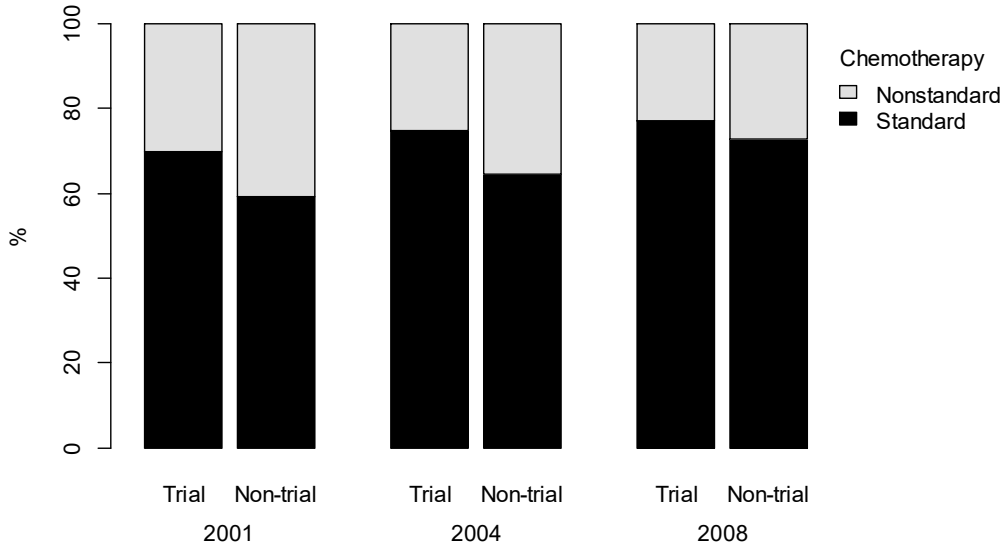

Figure 16: Chemotherapy in advanced-stage ovarian cancer in trial and non-trial hospitals by year of diagnosis

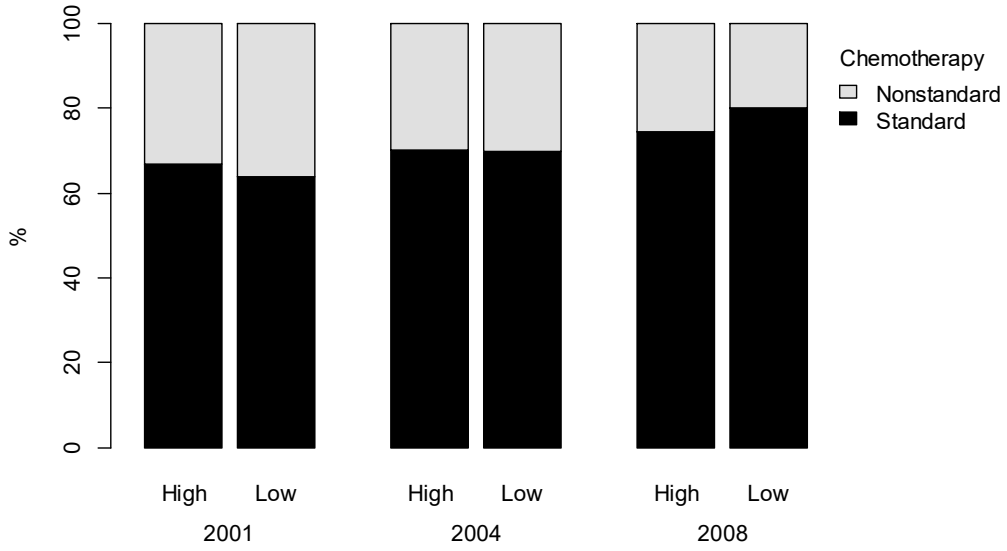

Figure 17: Chemotherapy in advanced-stage ovarian cancer in high-volume (12+ patients/year) and low-volume (1–11 patients/year) hospitals by year of diagnosis

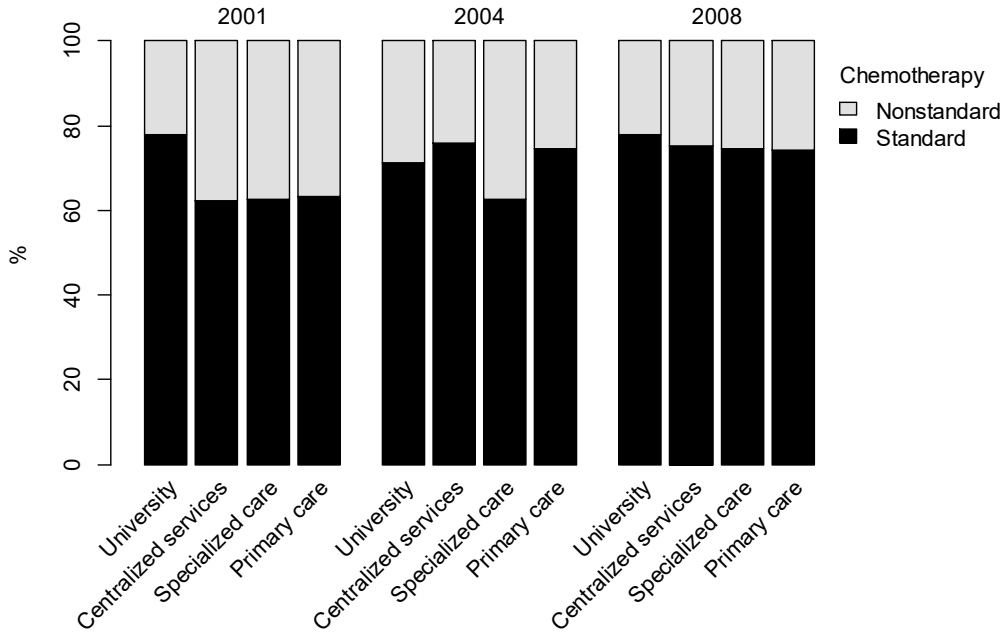

Figure 18: Chemotherapy in advanced-stage ovarian cancer in university hospitals, hospitals with centralized services and maximum care, hospitals providing specialized care and primary care hospitals by year of diagnosis

3.3.2.3 Model for adherence to chemotherapy guidelines

In the next step, the effect of hospital research activity on the appropriateness of chemotherapy was explored in more detail. Adjuvant platinum-based chemotherapy was regarded standard care for patients with FIGO stage I–IIA ovarian cancer, except for patients with FIGO IA G1 tumors who should not receive chemotherapy at all. For advanced-stage ovarian cancer, any combination containing carboplatin and paclitaxel was considered chemotherapy standard. Logistic regression models were fitted with generalized estimating equations (GEE) to account for clustering of patients within hospitals. Again, patients of all stages of disease were pooled and the predictors of adherence to treatment guidelines were determined for each of the three QS-OVAR cohorts. The same baseline covariates were used as for modeling adherence to surgical guidelines: stage of disease (FIGO I–IIA vs. FIGO IIB–IV), patient age at diagnosis (continuous, in units of 5 years), ECOG performance status (> 1 vs. 0/1), clinically relevant amount of ascites (> 500 ml vs. ≤ 500 ml), comorbidity (present vs. none), histology (serous vs. other), and grade (G3/4 vs. G1/2).

Table 7 shows the impact of institutional research activity on the patient's likelihood of receiving guideline-adherent chemotherapy after adjustment for potential baseline confounders. Treatment in a trial hospital seemed to be associated with standard chemotherapy when compared to treatment in a non-trial hospital in two of the three cohorts. In QS-OVAR 2004, the adjusted odds ratio in favor of trial hospitals was 1.65 (95% CI 1.11 to 2.45). Even though statistically non-significant, a similar result was observed in the first cohort (OR = 1.58, 95% CI 0.99 to 2.52). In contrast, no positive effect of institutional research activity was visible in QS-OVAR 2008 (OR = 0.93, 95% CI 0.65 to 1.33). Similar to the pattern observed for surgery, and consistent among all three QS-OVAR cohorts, older age and worse performance status were found to be predictors of receiving suboptimal chemotherapy treatment. In all three cohorts, women with comorbid conditions were significantly less likely to be optimally treated with chemotherapy compared to women without comorbidities.

Table 6A: Chemotherapy in FIGO IA G1 ovarian cancer

Chemotherapy (CT)	QS-OVAR 2001 hospitals			QS-OVAR 2004 hospitals			QS-OVAR 2008 hospitals		
	Trial (*N* = 8)	Non-trial (*N* = 12)	All (*N* = 20)	Trial (*N* = 11)	Non-trial (*N* = 9)	All (*N* = 20)	Trial (*N* = 7)	Non-trial (*N* = 3)	All (*N* = 10)
No CT (standard), *N* (%)	7 (87.5)	11 (91.7)	18 (95.0)	10 (90.9)	7 (77.8)	17 (85.0)	7 (100.0)	3 (100.0)	10 (100.0)
CT (overtreatment)	1 (12.5)	1 (8.3)	2 (5.0)	1 (9.1)	2 (22.2)	3 (15.0)	0 (0.0)	0 (0.0)	0 (0.0)

Table 6B: Chemotherapy in FIGO IA G2+ and IB–IIA ovarian cancer

Chemotherapy (CT)	QS-OVAR 2001 hospitals			QS-OVAR 2004 hospitals			QS-OVAR 2008 hospitals		
	Trial (*N* = 48)	Non-trial (*N* = 56)	All (*N* = 104)	Trial (*N* = 84)	Non-trial (*N* = 79)	All (*N* = 163)	Trial (*N* = 114)	Non-trial (*N* = 57)	All (*N* = 171)
Pt-based CT (standard), *N* (%)	33 (68.8)	33 (58.9)	66 (63.5)	59 (70.2)	51 (64.6)	110 (67.5)	89 (78.1)	45 (78.9)	134 (78.4)
Other CT (substandard)	2 (4.2)	0 (0.0)	2 (1.9)	1 (1.2)	3 (3.8)	4 (2.5)	0 (0.0)	1 (1.8)	1 (0.6)
No CT (nonstandard)	13 (27.1)	23 (41.1)	36 (34.6)	24 (28.6)	25 (31.6)	49 (30.1)	25 (21.9)	11 (19.3)	36 (21.1)

Table 6C: Chemotherapy in FIGO IIB–IV ovarian cancer

Chemotherapy (CT)	QS-OVAR 2001 hospitals			QS-OVAR 2004 hospitals			QS-OVAR 2008 hospitals		
	Trial (*N* = 219)	Non-trial (*N* = 133)	All (*N* = 352)	Trial (*N* = 318)	Non-trial (*N* = 262)	All (*N* = 580)	Trial (*N* = 449)	Non-trial (*N* = 251)	All (*N* = 700)
Pt + Taxan (standard), *N* (%)	153 (69.9)	79 (59.4)	232 (65.9)	238 (74.8)	169 (64.5)	407 (70.2)	346 (77.1)	183 (72.9)	529 (75.6)
Pt without Taxan	23 (10.5)	24 (18.0)	47 (13.4)	32 (10.1)	34 (13.0)	66 (11.4)	30 (6.7)	35 (13.9)	65 (9.3)
CT without Pt	12 (5.5)	8 (6.0)	20 (5.7)	8 (2.5)	10 (3.8)	18 (3.1)	11 (2.4)	6 (2.4)	17 (2.4)
No CT	31 (14.2)	22 (16.5)	53 (15.1)	40 (12.6)	49 (18.7)	89 (15.3)	62 (13.8)	27 (10.8)	89 (12.7)

Note: Pt = Platinum

Table 7: Predictors of optimal chemotherapy*

	QS-OVAR 2001					QS-OVAR 2004					QS-OVAR 2008				
	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI
Research Activity															
Non-trial hospital	201	123	(61.2)	1	Reference	350	227	(64.9)	1	Reference	311	231	(74.3)	1	Reference
Trial hospital	275	193	(70.2)	1.58	0.99 to 2.52	413	307	(74.3)	1.65	1.11 to 2.45	570	442	(77.5)	0.93	0.65 to 1.33
FIGO stage															
I–IIA	124	84	(67.7)	1	Reference	183	127	(69.4)	1	Reference	181	144	(79.6)	1	Reference
IIB–IV	352	232	(65.9)	1.01	0.56 to 1.84	580	407	(70.2)	1.29	0.80 to 2.08	700	529	(75.6)	0.73	0.43 to 1.23
Age (continuous)															
per 5 years	476	316	(66.4)	0.75	0.66 to 0.85	763	534	(70.0)	0.75	0.68 to 0.83	881	673	(76.4)	0.69	0.61 to 0.78
ECOG performance															
0/1	374	284	(75.9)	1	Reference	564	448	(79.4)	1	Reference	657	558	(84.9)	1	Reference
> 1	102	32	(31.4)	0.28	0.16 to 0.50	199	86	(43.2)	0.37	0.24 to 0.57	224	115	(51.3)	0.35	0.23 to 0.53
Ascites															
≤ 500 ml	284	189	(66.5)	1	Reference	413	292	(70.7)	1	Reference	455	352	(77.4)	1	Reference
> 500 ml	192	127	(66.1)	1.28	0.81 to 2.01	350	242	(69.1)	1.13	0.75 to 1.69	426	321	(75.4)	1.35	0.95 to 1.92
Comorbidity															
None	360	267	(74.2)	1	Reference	561	438	(78.1)	1	Reference	587	497	(84.7)	1	Reference
Present	116	49	(42.2)	0.55	0.34 to 0.87	202	96	(47.5)	0.45	0.29 to 0.71	294	176	(59.9)	0.51	0.36 to 0.73
Histology															
Other	147	92	(62.6)	1	Reference	282	185	(65.6)	1	Reference	265	200	(75.5)	1	Reference
Serous	329	224	(68.1)	1.38	0.84 to 2.28	481	349	(72.6)	1.27	0.89 to 1.83	616	473	(76.8)	1.22	0.78 to 1.90
Grade															
G1/G2	261	169	(64.8)	1	Reference	409	278	(68.0)	1	Reference	382	284	(74.3)	1	Reference
G3/G4	215	147	(68.4)	1.39	0.88 to 2.19	354	256	(72.3)	1.59	1.09 to 2.32	499	389	(78.0)	1.31	0.89 to 1.93

Note: *Optimal chemotherapy: Platinum-based (FIGO I–IIA, except for FIGO IA G1)/Platinum-taxane (FIGO IIB–IV)

3.4 Survival

The survival information was collected yearly for at least three years after diagnosis in QS-OVAR 2001. In the next two cohorts, the minimum follow-up was extended to four years. Consistent with standard practice in survival studies, the minimum follow-up time refers to the last patient included. The reverse Kaplan-Meier estimates of median follow-up were 35.1 months for the first cohort, 51.7 months for the second cohort, and 48.5 months for the last cohort. During these follow-up periods, a total of 199 patients out of 476 died in QS-OVAR 2001, 395 of 763 patients died in QS-OVAR 2004, and 451 of 881 patients died in QS-OVAR 2008. Figure 19 shows the Kaplan-Meier estimates of survival in the three cohorts. The estimated median survival times were 37.1 months in QS-OVAR 2001, 44.8 months in QS-OVAR 2004, and 43.5 months in QS-OVAR 2008. The 1-year survival rates were similar among all three cohorts: 82% in the first cohort, 81% in the second cohort and 83% in the third cohort. The corresponding 3-year survival rates were 54%, 56%, and 57%, respectively.

During the follow-up, only 15 deaths occurred in the 124 patients with early ovarian cancer in the first cohort. Similarly, only 24 of the 183 patients with FIGO stage I–IIA died in the second cohort. Finally, 41 of the 181 patients with FIGO stage I–IIA died in QS-OVAR 2008. Consequently, for early-stage ovarian cancer, the median was not reached at the time of analysis in any of the three cohorts. Instead, most deaths occurred in patients with advanced ovarian cancer: In the first cohort, 184 of the 352 patients with FIGO stage IIB–IV died, in the second cohort these were 371 of the 580 patients, and in the last cohort 410 of the 700 patients with FIGO stage IIB–IV died. The median survival times for patients with FIGO stage IIB or higher were 31.0 months in QS-OVAR 2001, 29.7 months in QS-OVAR 2004, and 36.2 months in QS-OVAR 2008. The 1-year survival rates were 78% in the first cohort, 76% in the second cohort and 80% in the third cohort, respectively. The 3-year survival rates were 43% in QS-OVAR 2001, 45% in QS-OVAR 2004, and 50% in QS-OVAR 2008, respectively.

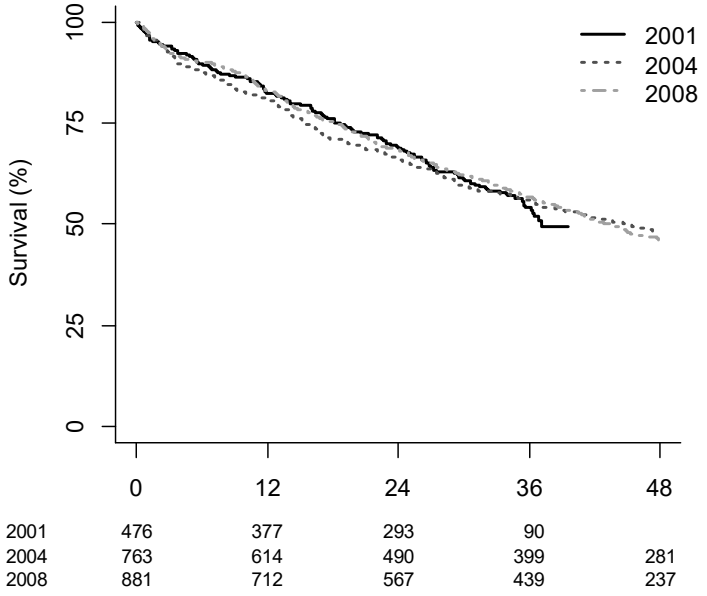

Figure 19: Survival in the three QS-OVAR cohorts (time axis in months). The numbers of patients at risk is shown below the time axis.

3.4.1 Adherence to treatment guidelines and survival

One main question of the thesis was whether the hypothesized effect of hospital trial participation on patient survival is mediated through better adherence to treatment guidelines with regard to surgery and chemotherapy. Therefore, the impact of the two variables representing adherence to surgical and chemotherapy treatment guidelines on survival was investigated in more detail.

When all three cohorts were taken together for descriptive analysis, 506 deaths were observed among the 1230 patients optimally treated surgically and 539 deaths were observed among the 890 patients who were not treated in accordance to surgical guidelines; the latter meaning that patients were either not optimally staged or had tumor residuals of more than 10 mm after surgery. Optimally treated patients had considerably longer survival than patients who did not receive the optimal surgical treatment (median 52.9 months vs. 27.1 months). In the first two cohorts, the median survival time for optimally treated patients was not reached at the time of analysis, whereas the median survival time in non-optimally treated patients was 34.7 months in QS-OVAR 2001 and 25.9 months in QS-OVAR 2004. In QS-OVAR 2008, the median survival time was 52.3 months for optimally treated patients and 22.2 months for patients who received suboptimal surgical treatment. Figure 20 shows the corresponding Kaplan-Meier survival curves by stage of disease and cohort. As can be seen from this figure, the pattern was quite consistent for patients with early and advanced ovarian cancer in all three cohorts with higher survival rates for patients treated according to surgical guidelines.

Similar results were observed with regard to adherence to chemotherapy guidelines. In the pooled cohorts, 642 deaths were observed among the 1523 patients who were treated according to chemotherapy guidelines and 403 deaths were observed among the 597 patients with non-optimal chemotherapy; the latter including those few patients without chemotherapy despite treatment recommendations. As with surgery, patients treated according to chemotherapy guidelines had considerably longer survival than patients who did not receive the appropriate chemotherapy (median 52.3 months vs. 18.1 months). This pattern was consistent in all three cohorts. In the first two cohorts, the median survival time for optimally treated patients was not reached, whereas the median survival time in non-optimally treated patients was 22.3 months in QS-OVAR 2001 and 16.9 months in QS-OVAR 2004. In the third cohort, the median survival time was 50.5 months for optimally treated patients and 15.6 months for patients who did not receive the optimal chemotherapeutic treatment. The picture remained the same when patients with early and advanced ovarian cancer were considered separately (Figure 21).

Multivariable Cox models confirmed the findings described above. In these models, relevant patient and disease characteristics as well as clustering of patients into hospitals were controlled for. Guideline adherence to both surgery and chemotherapy showed substantial beneficial effect on patient survival (Table 8). In all three QS-OVAR cohorts, the risk of death was reduced by around 50% when patients were treated in accordance to the surgical guidelines. Similarly, patients treated optimally with regard to chemotherapy had a risk reduction of about 50% compared to patients who were not optimally treated chemotherapeutically.

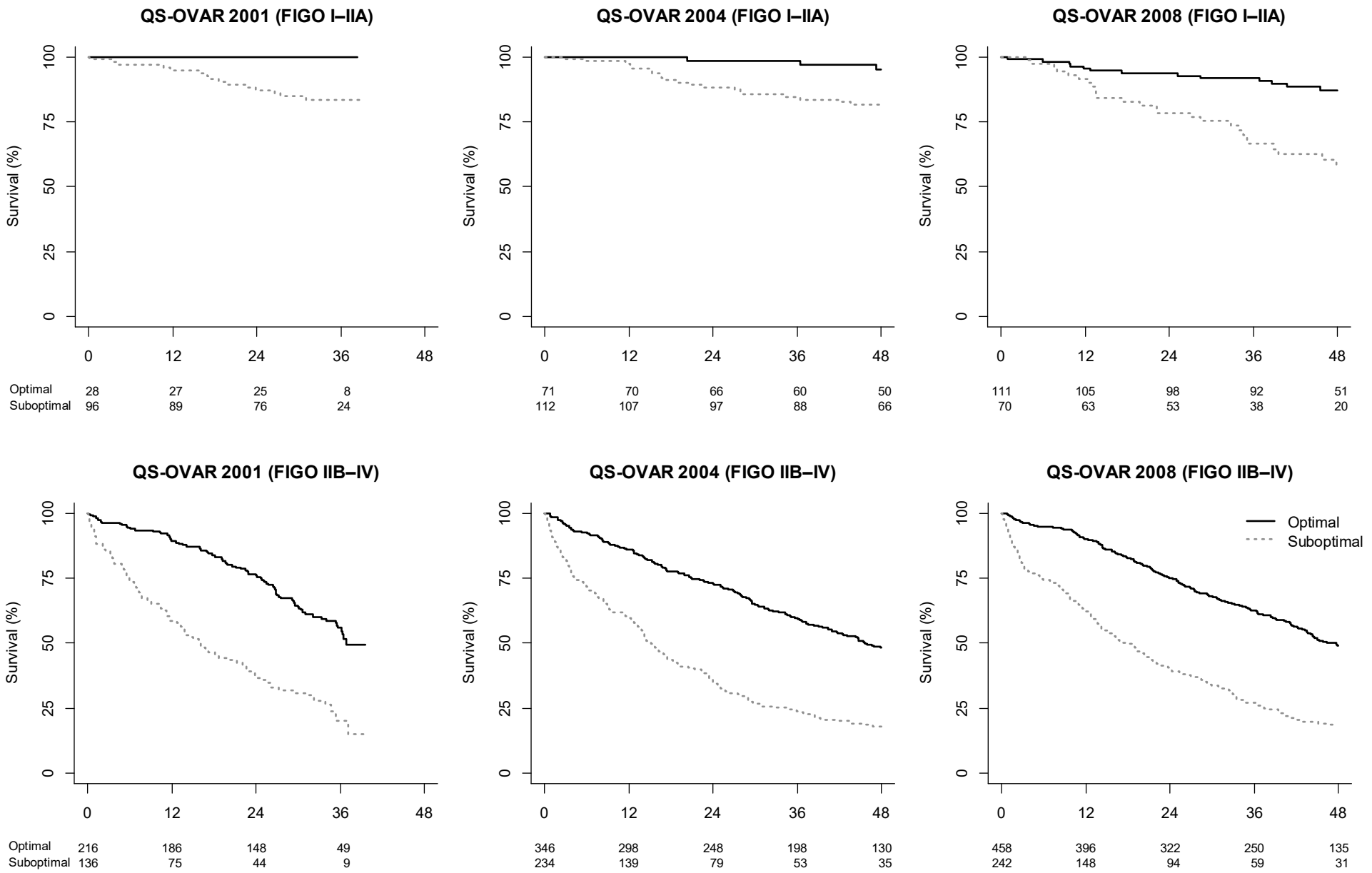

Figure 20: The relationship of patient survival and optimality of surgery (for early-stage and late-stage ovarian cancer)

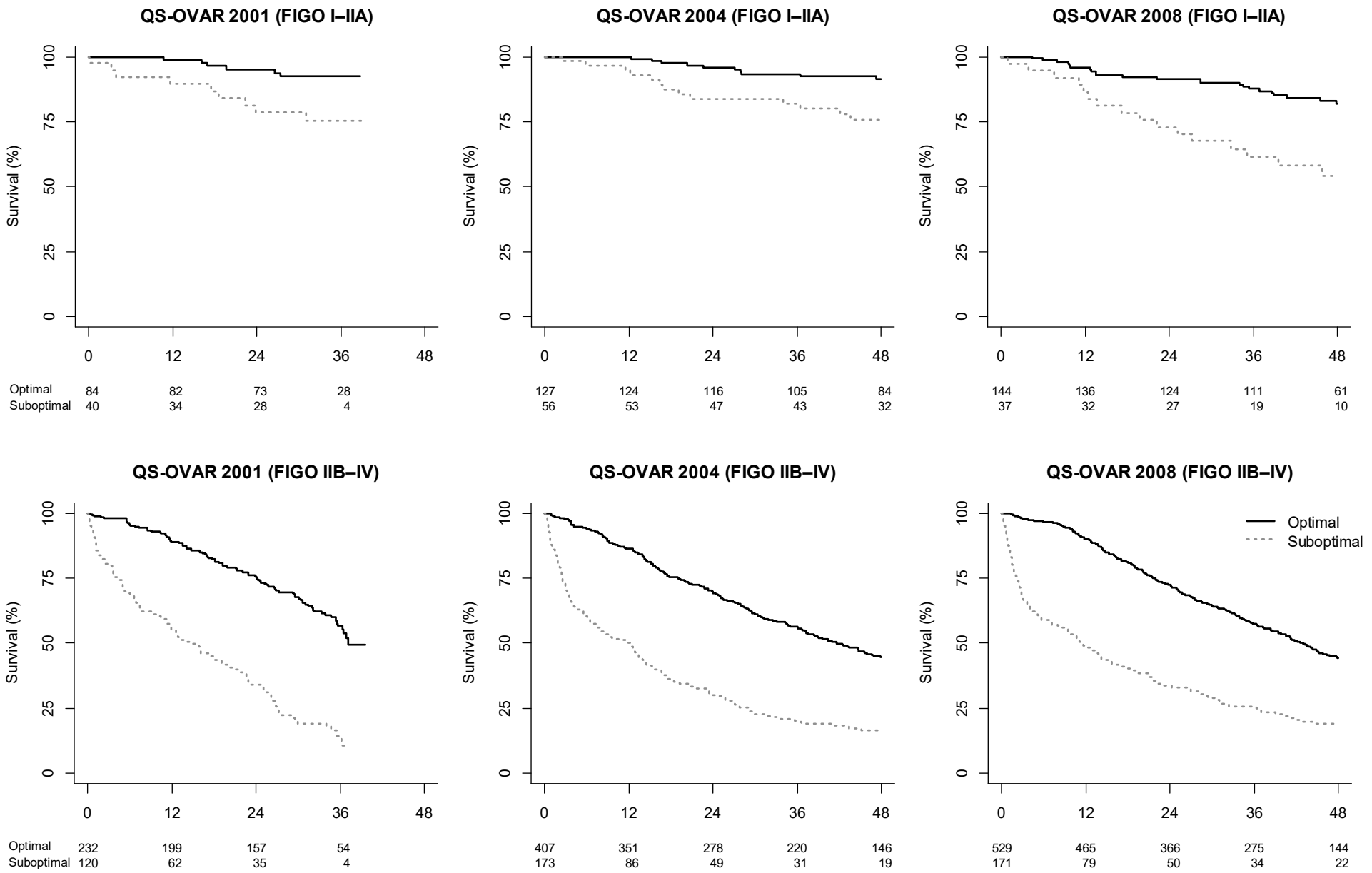

Figure 21: The relationship of patient survival and optimality of the chemotherapy (for early-stage and late-stage ovarian cancer)

Table 8: Survival depending on guideline adherence

	QS-OVAR 2001				QS-OVAR 2004				QS-OVAR 2008			
	Total	Deaths	HR	95% CI	Total	Deaths	HR	95% CI	Total	Deaths	HR	95% CI
*Optimal surgery**												
No	232	111	1	Reference	346	210	1	Reference	312	218	1	Reference
Yes	244	88	0.45	0.34 to 0.60	417	185	0.55	0.43 to 0.69	569	233	0.53	0.43 to 0.66
*Optimal chemotherapy***												
No	160	102	1	Reference	229	156	1	Reference	208	145	1	Reference
Yes	316	97	0.44	0.30 to 0.64	534	239	0.53	0.40 to 0.70	673	306	0.48	0.36 to 0.65
FIGO stage												
I–IIA	124	15	1	Reference	183	24	1	Reference	181	41	1	Reference
IIB–IV	352	184	7.28	3.91 to 13.56	580	371	7.46	4.62 to 12.03	700	410	2.90	2.03 to 4.14
Age (continuous)												
per 5 years	476	199	1.13	1.04 to 1.22	763	395	1.07	1.02 to 1.13	881	451	1.02	0.97 to 1.07
ECOG performance												
0/1	374	126	1	Reference	564	239	1	Reference	657	286	1	Reference
> 1	102	73	1.55	1.05 to 2.28	199	156	1.58	1.21 to 2.06	224	165	1.49	1.18 to 1.88
Ascites												
≤ 500 ml	284	88	1	Reference	413	142	1	Reference	455	160	1	Reference
> 500 ml	192	111	1.51	1.07 to 2.14	350	253	1.83	1.43 to 2.34	426	291	1.87	1.47 to 2.37
Comorbidity												
None	360	122	1	Reference	561	250	1	Reference	587	269	1	Reference
Present	116	77	1.41	1.02 to 1.97	202	145	1.58	1.22 to 2.06	294	182	1.27	1.02 to 1.59
Histology												
Other	147	47	1	Reference	282	138	1	Reference	265	109	1	Reference
Serous	329	152	1.24	0.85 to 1.79	481	257	0.87	0.67 to 1.11	616	342	0.97	0.76 to 1.24
Grade												
G1/G2	261	89	1	Reference	409	189	1	Reference	382	168	1	Reference
G3/G4	215	110	1.05	0.77 to 1.42	354	206	0.92	0.75 to 1.14	499	283	1.04	0.85 to 1.25

Note: *max. 1 missing staging item (FIGO I–IIA)/max. 10 mm tumor residual (FIGO IIB–IV); **Platinum-based (FIGO I–IIA, except for FIGO IA G1)/Platinum-taxane (FIGO IIB–IV)

3.4.2 Hospital research activity and survival

Overall, patients in trial hospitals had longer survival than patients treated in non-trial hospitals (median 45.0 months vs. 41.6 months) with 601 deaths observed among the 1258 patients in trial hospitals compared to 444 deaths observed among the 862 patients in non-trial hospitals. However, the picture was inconsistent across cohorts. In QS-OVAR 2001, the median survival was not reached for patients treated in trial hospitals, whereas it was 37.1 months for patients in non-trial hospitals. Similarly, in QS-OVAR 2008, the median survival time was longer for patients treated in trial hospitals compared to patients treated in non-trial hospitals (47.7 months vs. 38.9 months). In contrast, in QS-OVAR 2004, the median survival time for patients in trial hospitals was even shorter than that for patients treated in non-trial hospitals (40.7 vs. 46.9 months). Figure 22 shows the Kaplan-Meier survival curves for patients in trial hospitals versus patients in non-trial hospitals by stage of disease and QS-OVAR cohort. Here, a clear beneficial effect of institutional trial participation on survival was observed only in patients with advanced ovarian cancer in QS-OVAR 2001 (bottom left panel).

Multivariable analysis confirmed the findings described above. Table 9 shows the predictors of survival including institutional research activity as well as the known relevant patient and disease characteristics. It presents the total number of patients for each predictor category, the number of deaths, and hazard ratios with their 95% confidence intervals from multivariable Cox regression models that were adjusted for clustering of patients into hospitals. In QS-OVAR 2001, the patients treated in trial hospitals had significantly longer survival than those patients treated in non-trial hospitals (hazard ratio 0.56 with 95% CI 0.42 to 0.76). However, in the following two QS-OVAR cohorts, no such beneficial effect of hospital trial participation was detected. The corresponding hazard ratios for trial hospitals versus non-trial hospitals were 1.18 (95% CI 0.93 to 1.50) in QS-OVAR 2004 and 0.97 (95% CI 0.80 to 1.17) in QS-OVAR 2008, respectively.

Patients with FIGO Stage IIB–IV had a higher risk of death than patients with FIGO Stage I–IIA. Similarly, older patients, patients with worse ECOG performance status, ascites over 500 ml, and comorbidity had a worse prognosis. No consistent associations were found for histological subtype and grade of the tumor. Compared to the analysis shown in Table 8 that investigated the effect of adherence to treatment guidelines on patient survival instead of hospital research activity, the prognostic value of the included patient and disease characteristics remained mostly unchanged.

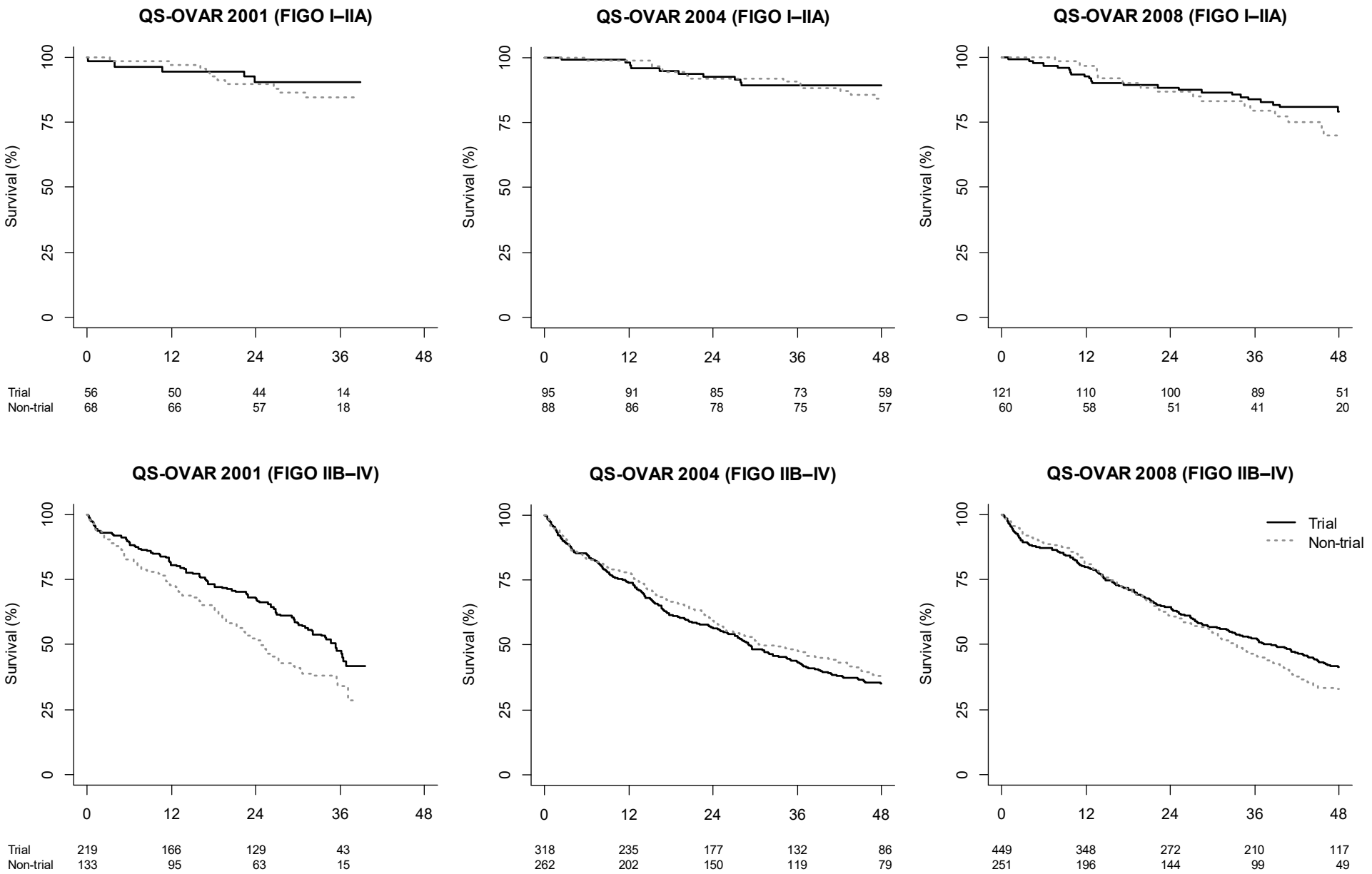

Figure 22: The relationship of patient survival and research activity in the three QS-OVAR cohorts (for early-stage and late-stage ovarian cancer)

Table 9: Hospital research activity and survival

	QS-OVAR 2001				QS-OVAR 2004				QS-OVAR 2008			
	Total	Deaths	HR	95% CI	Total	Deaths	HR	95% CI	Total	Deaths	HR	95% CI
Research activity												
Non-trial hospital	201	91	1	Reference	350	177	1	Reference	311	176	1	Reference
Trial hospital	275	108	0.56	0.42 to 0.76	413	218	1.18	0.93 to 1.50	570	275	0.97	0.80 to 1.17
FIGO stage												
I–IIA	124	15	1	Reference	183	24	1	Reference	181	41	1	Reference
IIB–IV	352	184	5.04	2.79 to 9.09	580	371	5.28	3.36 to 8.29	700	410	2.38	1.70 to 3.32
Age (continuous)												
per 5 years	476	199	1.22	1.13 to 1.32	763	395	1.15	1.10 to 1.22	881	451	1.11	1.06 to 1.16
ECOG performance												
0/1	374	126	1	Reference	564	239	1	Reference	657	286	1	Reference
> 1	102	73	2.07	1.44 to 2.99	199	156	1.86	1.41 to 2.45	224	165	1.78	1.41 to 2.23
Ascites												
≤ 500 ml	284	88	1	Reference	413	142	1	Reference	455	160	1	Reference
> 500 ml	192	111	1.69	1.19 to 2.38	350	253	1.95	1.54 to 2.47	426	291	2.00	1.60 to 2.49
Comorbidity												
None	360	122	1	Reference	561	250	1	Reference	587	269	1	Reference
Present	116	77	1.58	1.18 to 2.11	202	145	1.69	1.29 to 2.20	294	182	1.43	1.15 to 1.77
Histology												
Other	147	47	1	Reference	282	138	1	Reference	265	109	1	Reference
Serous	329	152	1.24	0.85 to 1.81	481	257	0.86	0.68 to 1.10	616	342	0.95	0.74 to 1.23
Grade												
G1/G2	261	89	1	Reference	409	189	1	Reference	382	168	1	Reference
G3/G4	215	110	1.13	0.83 to 1.54	354	206	0.91	0.74 to 1.12	499	283	1.05	0.87 to 1.27

3.4.3 Mediation analysis

One major goal of this thesis was to introduce a new methodology that can be applied to assess mediation in survival data, and to use this method to investigate the effects of institutional research activity on survival of ovarian cancer patients in presence of two potential mediators, namely the guideline adherent surgery and chemotherapy as defined in Section 2.1. The impact of hospital participation on overall survival in terms of natural direct and indirect effects can be summarized as follows (Table 10): In QS-OVAR 2001, a total hazard ratio of 0.57 in favor of institutional participation in clinical trials was observed. This total effect was decomposed into a direct hazard ratio for research activity of 0.66 (95% CI 0.48 to 0.91) and an indirect hazard ratio for the mediators of 0.86 (95% CI 0.76 to 0.96). The indirect effect corresponding to surgery alone was 0.91, the indirect effect corresponding to chemotherapy alone was 0.94; thus, the resulting total effect was $0.66 \times 0.91 \times 0.94 = 0.57$. The proportion mediated through surgery and chemotherapy was 17% and 11% on the log scale, respectively. Taken together, about 27% of the total effect of hospital research activity on survival was mediated through the proposed pathways (95% CI 8% to 67%). In the following two cohorts, the indirect effects were still present but decreased in magnitude: The indirect effect via the two mediators together was 0.90 in QS-OVAR 2004 and 0.96 in QS-OVAR 2008, respectively. Surprisingly, the cohorts differed regarding the direct effects. In particular, the direct effect of 1.32 in QS-OVAR 2004 was not in favor of research active institutions, but in favor of non-trial hospitals. The direct effect in QS-OVAR 2008 was 1.01. Besides that, the mediation models were quite stable and the prognostic value of the included patient and disease characteristics remained unchanged compared to Cox models (see Table 9).

Table 10: Mediation analysis

	QS-OVAR 2001		QS-OVAR 2004		QS-OVAR 2008	
Effect	HR	95% CI	HR	95% CI	HR	95% CI
Trial hospital						
Total effect	0.57	0.42 to 0.77	1.18	0.93 to 1.53	0.97	0.79 to 1.17
Direct effect	0.66	0.48 to 0.91	1.32	1.04 to 1.70	1.01	0.83 to 1.21
Indirect effect	0.86	0.76 to 0.96	0.90	0.83 to 0.96	0.96	0.90 to 1.03
via Surgery	0.91	0.83 to 0.99	0.95	0.90 to 1.00	0.95	0.90 to 1.00
via Chemotherapy	0.94	0.87 to 1.00	0.94	0.89 to 0.99	1.01	0.97 to 1.05
Covariates						
Age (per 5 years)	1.22	1.13 to 1.32	1.15	1.10 to 1.21	1.11	1.06 to 1.16
Stage (IIB–IV)	5.04	2.79 to 9.09	5.29	3.36 to 8.29	2.38	1.71 to 3.32
ECOG (> 1)	2.08	1.44 to 2.99	1.86	1.41 to 2.45	1.78	1.41 to 2.23
Ascites (> 500 ml)	1.69	1.20 to 2.38	1.95	1.54 to 2.47	2.00	1.60 to 2.49
Comorbidity (present)	1.58	1.18 to 2.11	1.69	1.29 to 2.20	1.43	1.15 to 1.77
Histology (serous)	1.24	0.85 to 1.82	0.86	0.68 to 1.10	0.95	0.74 to 1.23
Grade (3/4)	1.13	0.83 to 1.54	0.91	0.74 to 1.12	1.05	0.87 to 1.26

Note: Minor numeric inconsistencies between Table 9 and Table 10 are a consequence of the effect decomposition that includes reweighting and a quadruplication of the data.

3.5 Sensitivity analyses

3.5.1 Hospital volume and survival

One additional question of the present thesis was whether other structural measures than institutional participation in clinical trials were potentially associated with longer patient survival. In particular, hospital volume was of interest as volume-outcome relationship has been found for surgical procedures in cancer patients and other serious diseases (see Section 1.2 for more details). Hospital volume was considered as a continuous variable with 1 patient per year but also as a dichotomous variable with low-volume hospitals treating 1 to 11 patients per year and high-volume hospitals treating 12 or more patients per year, respectively.

Figure 23 shows the survival data for patients with early and advanced ovarian cancer treated in the two hospital volume categories. Examination of the Kaplan-Meier plots suggested an association towards longer survival in high-volume hospitals in patients with advanced ovarian cancer in the first cohort. However, overall, there was no clear picture of high hospital volume resulting in a survival benefit among the three cohorts of patients with ovarian cancer.

In order to assess the impact of hospital volume on patient survival while adjusting for other covariates at hospital and patient level and to determine whether the previous results (Table 9) might have been influenced by hospital volume, the Cox regression analyses were repeated with this additional covariate. As can be seen from Table 11, with the inclusion of hospital volume the results for hospital research activity remained essentially unchanged. Similarly, the estimates for the patient and disease characteristics did not considerably differ from previous estimates after including hospital volume into the models (not shown). This held independent of whether the continuous or categorical hospital volume was added to the regression models.

3.5.2 Hospital care level and survival

Finally, the association between survival and another structural measure, namely the hospital level of care, was investigated in more detail to get a complete picture with regard to the collected hospital characteristics in QS-OVAR.

Figure 24 shows the Kaplan-Meier plots for patients with early and advanced ovarian cancer treated in the four categories of hospitals (i.e., university hospitals, hospitals with centralized services and maximum care, hospitals providing specialized care, and primary care hospitals). The figure does not reveal clear associations between hospital care level and patient survival. In line with this, the results for hospital research activity and hospital volume remained practically the same after hospital care level had been included as additional factor in the models that already contained the two other hospital characteristics (Table 12). Similarly, the estimates for the patient and disease characteristics did not considerably change from previous estimates (not shown).

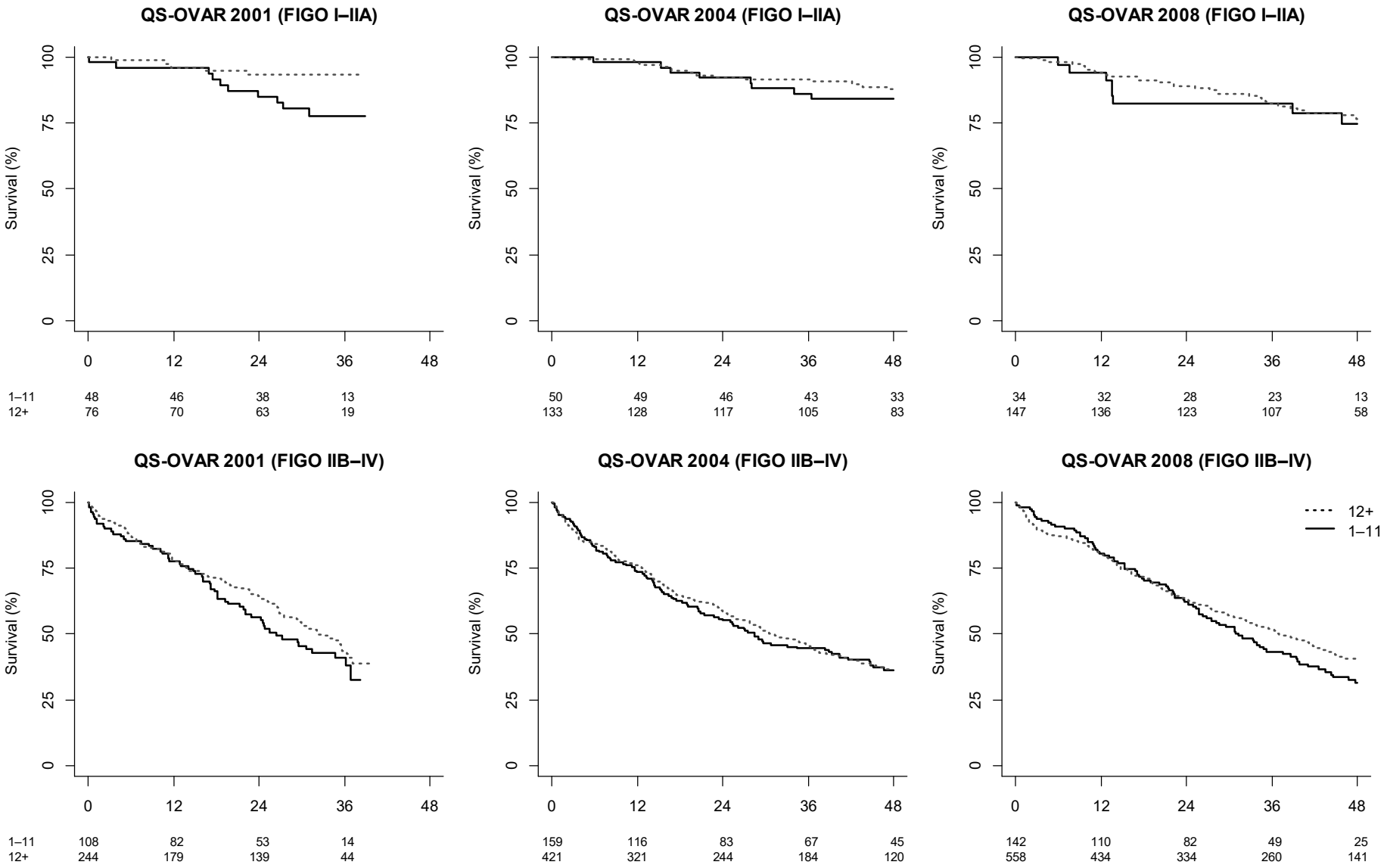

Figure 23: The relationship of patient survival and hospital volume in the three QS-OVAR cohorts (for early-stage and late-stage ovarian cancer)

Table 11A: Hospital research activity, hospital volume and survival

	QS-OVAR 2001				QS-OVAR 2004				QS-OVAR 2008			
	To-tal	Deaths	HR	95% CI	To-tal	Deaths	HR	95% CI	To-tal	Deaths	HR	95% CI
Research activity												
Non-trial hospital	201	91	1	Reference	350	177	1	Reference	311	176	1	Reference
Trial hospital	275	108	0.58	0.42 to 0.79	413	218	1.14	0.88 to 1.48	570	275	0.99	0.80 to 1.23
Hospital volume (continuous covariate)												
per 1 pt/year	476	199	0.998	0.991 to 1.005	763	395	1.004	0.998 to 1.011	881	451	0.999	0.994 to 1.003

Table 11B: Hospital research activity, hospital volume and survival

Research activity												
Non-trial hospital	201	91	1	Reference	350	177	1	Reference	311	176	1	Reference
Trial hospital	275	108	0.58	0.42 to 0.80	413	218	1.16	0.90 to 1.48	570	275	0.98	0.80 to 1.21
Hospital volume (categorical covariate)												
1–11 pts/year	156	70	1	Reference	209	108	1	Reference	176	97	1	Reference
12+ pts/year	320	129	0.90	0.66 to 1.23	554	287	1.16	0.88 to 1.52	705	354	0.95	0.76 to 1.20

Note: Models include stage of disease (FIGO I–IIA vs. IIB–IV), age at diagnosis (continuous, 5 years), ECOG performance status (>1 vs. 0/1), ascites (>500 ml vs. ≤500 ml), comorbidity (present vs. none), tumor histology (serous vs. other), and tumor grade (G3/4 vs. G1/2) and take clustering of patients in hospitals into account.

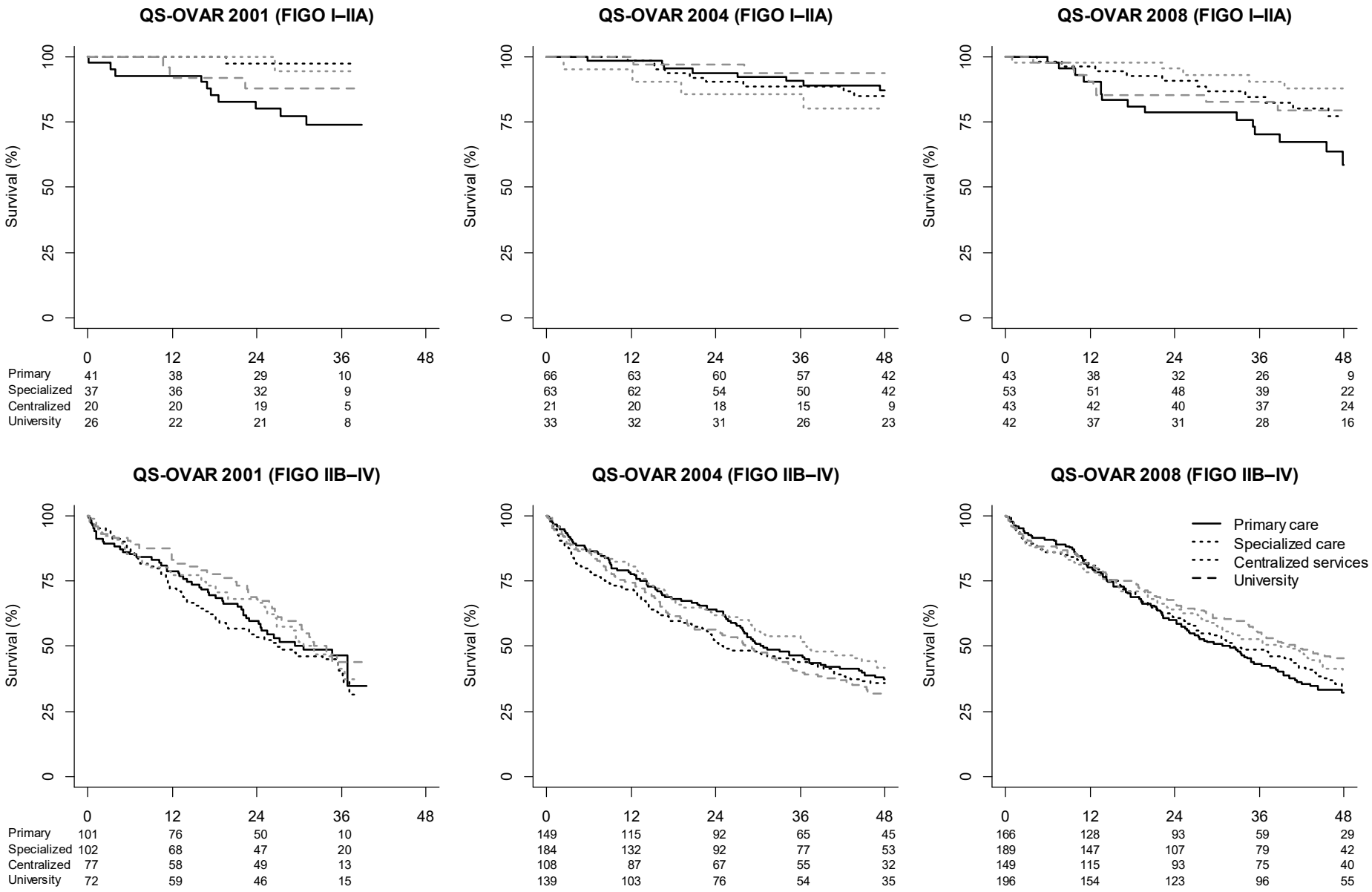

Figure 24: The relationship of patient survival and hospital care level in the three QS-OVAR cohorts (for early-stage and late-stage ovarian cancer)

Table 12A: Hospital research activity, hospital volume, hospital care level and survival

	QS-OVAR 2001				QS-OVAR 2004				QS-OVAR 2008			
	To-tal	Deaths	HR	95% CI	To-tal	Deaths	HR	95% CI	To-tal	Deaths	HR	95% CI
Research activity												
Non-trial hospital	201	91	1	Reference	350	177	1	Reference	311	176	1	Reference
Trial hospital	275	108	0.57	0.40 to 0.83	413	218	1.14	0.86 to 1.49	570	275	1.02	0.82 to 1.27
Hospital volume (continuous covariate)												
per 1 pt/year	476	199	0.998	0.988 to 1.007	763	395	1.006	0.997 to 1.015	881	451	1.000	0.995 to 1.005
Hospital care level												
Primary care	142	59	1	Reference	215	112	1	Reference	209	119	1	Reference
Specialized care	139	58	0.99	0.65 to 1.52	247	130	0.97	0.73 to 1.28	242	126	0.94	0.74 to 1.18
Centralized services	97	42	0.91	0.57 to 1.43	129	67	0.80	0.47 to 1.37	192	97	0.88	0.65 to 1.18
University	98	40	1.03	0.54 to 1.95	172	96	1.01	0.71 to 1.43	238	109	0.83	0.59 to 1.15

Table 12B: Hospital research activity, hospital volume, hospital care level and survival

Research activity												
Non-trial hospital	201	91	1	Reference	350	177	1	Reference	311	176	1	Reference
Trial hospital	275	108	0.58	0.40 to 0.84	413	218	1.14	0.86 to 1.50	570	275	1.02	0.82 to 1.26
Hospital volume (categorical covariate)												
1–11 pts/year	156	70	1	Reference	209	108	1	Reference	176	97	1	Reference
12+ pts/year	320	129	0.92	0.61 to 1.38	554	287	1.17	0.86 to 1.58	705	354	1.01	0.80 to 1.28
Hospital care level												
Primary care	142	59	1	Reference	215	112	1	Reference	209	119	1	Reference
Specialized care	139	58	1.01	0.62 to 1.64	247	130	0.95	0.72 to 1.26	242	126	0.93	0.74 to 1.19
Centralized services	97	42	0.91	0.57 to 1.47	129	67	0.84	0.53 to 1.35	192	97	0.88	0.66 to 1.18
University	98	40	0.99	0.55 to 1.78	172	96	1.05	0.77 to 1.45	238	109	0.83	0.62 to 1.11

Note: Models include stage of disease (FIGO I–IIA vs. IIB–IV), age at diagnosis (continuous, 5 years), ECOG performance status (>1 vs. 0/1), ascites (>500 ml vs. ≤500 ml), comorbidity (present vs. none), tumor histology (serous vs. other), and tumor grade (G3/4 vs. G1/2) and take clustering of patients in hospitals into account.

3.5.3 Adherence to surgical guidelines—complete surgery

3.5.3.1 Model for adherence to surgical guidelines

"Complete staging" and "complete debulking" were considered in a sensitivity analysis to take into account the more recent guidelines that define the primary objective of staging as the completion of all nine staging steps and debulking surgery in FIGO IIB–IV as the complete removal of all visible disease. As stated above (Section 3.3.1.1), only 102 of 488 patients (21%) with early-stage ovarian cancer received complete surgical staging (6% in 2001, 19% in 2004, and 34% in 2008). As described above (Section 3.3.1.2), complete debulking was achieved in 38% of the 1632 patients with advanced ovarian cancer (33% in 2001, 37% in 2004, and 42% in 2008). Even though, patients in trial hospitals seemed to have a slightly higher chance of receiving complete staging than patients in non-trial hospitals, only negligible differences were found between trial and non-trial hospitals regarding complete debulking in advanced-stage disease.

This was confirmed by a multivariable analysis: As can be seen from Table 13, in none of the three QS-OVAR cohorts, the patients treated in trial hospitals were significantly more likely to receive complete surgical treatment than those patients who were treated in non-trial hospitals. The adjusted odds ratios were 1.03 (95% CI 0.65 to 1.65) in QS-OVAR 2001, 1.25 in QS-OVAR 2004 (95% CI 0.85 to 1.85), and 1.14 (95% CI 0.81 to 1.61) in QS-OVAR 2008.

3.5.3.2 Survival

When all three cohorts were taken together for descriptive analysis, 249 deaths were observed among the 729 patients with "complete" outcome of surgery and 796 deaths were observed among the 1391 patients not treated in accordance with surgical guidelines; the latter meaning that patients were either not completely staged or had tumor residuals after surgery. Completely staged or debulked patients had considerably longer survival than patients with incomplete surgical treatment (median not reached vs. 33.3 months). In all three cohorts, the median survival time for completely treated patients was not reached at the time of analysis, whereas the median survival time in non-completely treated patients was 35.4 months in QS-OVAR 2001, 32.5 months in QS-OVAR 2004, and 31.9 months in QS-OVAR 2008. Figure 25 shows the corresponding Kaplan-Meier survival curves by stage of disease and QS-OVAR cohort. As can be seen from this figure, the pattern was quite consistent for patients with early and advanced ovarian cancer in all three cohorts with higher survival rates for patients treated according to surgical guidelines.

Multivariable Cox models confirmed the findings described above. In these models, chemotherapy treatment, relevant patient and disease characteristics as well as clustering of patients into hospitals were controlled for. In the three cohorts, the estimated reduction in the risk of death associated with guideline-recommended complete surgery was 55% to 64% (Table 14).

Table 13: Predictors of complete surgery*

	QS-OVAR 2001					QS-OVAR 2004					QS-OVAR 2008				
	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI
Research Activity															
Non-trial hospital	201	45	(22.4)	1	Reference	350	106	(30.3)	1	Reference	311	111	(35.7)	1	Reference
Trial hospital	275	79	(28.7)	1.03	0.65 to 1.65	413	143	(34.6)	1.25	0.85 to 1.85	570	245	(43.0)	1.14	0.81 to 1.61
FIGO stage															
I–IIA	124	7	(5.6)	1	Reference	183	34	(18.6)	1	Reference	181	61	(33.7)	1	Reference
IIB–IV	352	117	(33.2)	21.73	8.49 to 55.66	580	215	(37.1)	6.37	3.67 to 11.06	700	295	(42.1)	3.86	2.35 to 6.33
Age (continuous)															
per 5 years	476	124	(26.1)	0.83	0.75 to 0.92	763	249	(32.6)	0.91	0.85 to 0.98	881	356	(40.4)	0.83	0.77 to 0.89
ECOG performance															
0/1	374	106	(28.3)	1	Reference	564	208	(36.9)	1	Reference	657	300	(45.7)	1	Reference
> 1	102	18	(17.6)	0.80	0.37 to 1.70	199	41	(20.6)	0.54	0.34 to 0.86	224	56	(25.0)	0.70	0.44 to 1.10
Ascites															
≤ 500 ml	284	92	(32.4)	1	Reference	413	167	(40.4)	1	Reference	455	246	(54.1)	1	Reference
> 500 ml	192	32	(16.7)	0.22	0.12 to 0.38	350	82	(23.4)	0.28	0.18 to 0.43	426	110	(25.8)	0.21	0.15 to 0.29
Comorbidity															
None	360	103	(28.6)	1	Reference	561	198	(35.3)	1	Reference	587	259	(44.1)	1	Reference
Present	116	21	(18.1)	0.66	0.37 to 1.19	202	51	(25.2)	0.71	0.46 to 1.08	294	97	(33.0)	0.84	0.56 to 1.24
Histology															
Other	147	41	(27.9)	1	Reference	282	97	(34.4)	1	Reference	265	108	(40.8)	1	Reference
Serous	329	83	(25.2)	0.61	0.38 to 0.97	481	152	(31.6)	0.64	0.44 to 0.93	616	248	(40.3)	0.96	0.69 to 1.34
Grade															
G1/G2	261	57	(21.8)	1	Reference	409	126	(30.8)	1	Reference	382	152	(39.8)	1	Reference
G3/G4	215	67	(31.2)	1.41	0.85 to 2.35	354	123	(34.7)	1.03	0.71 to 1.49	499	204	(40.9)	1.01	0.73 to 1.39

Note: *Complete surgery: no missing staging item (FIGO I–IIA)/complete debulking (FIGO IIB–IV)

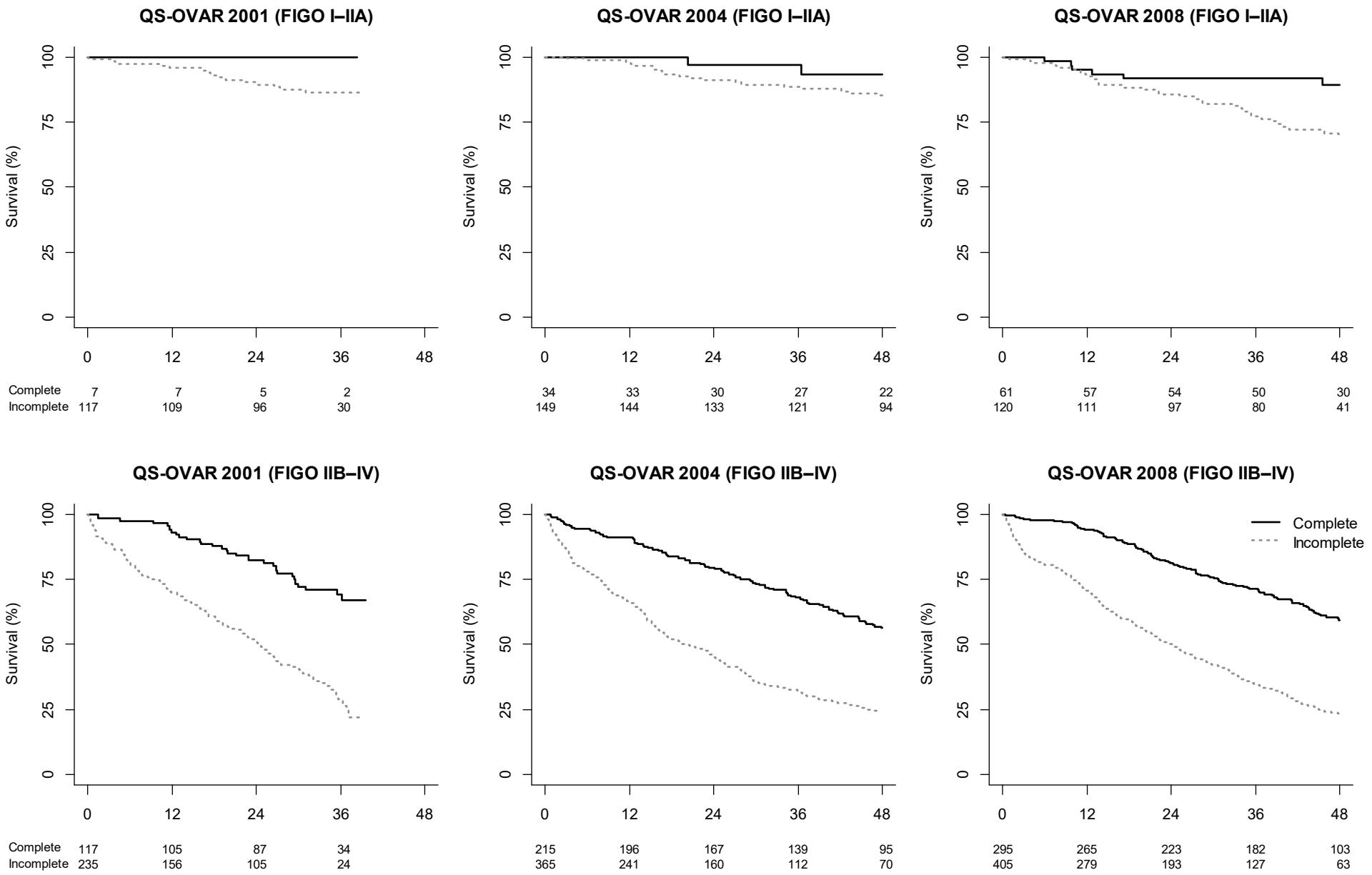

Figure 25: The relationship of patient survival and completeness of surgery (for early and late stage ovarian cancer)

Table 14: Predictors of survival

	QS-OVAR 2001				QS-OVAR 2004				QS-OVAR 2008			
	Total	Deaths	HR	95% CI	Total	Deaths	HR	95% CI	Total	Deaths	HR	95% CI
*Complete surgery**												
No	352	166	1	Reference	514	298	1	Reference	525	332	1	Reference
Yes	124	33	0.36	0.24 to 0.54	249	97	0.49	0.38 to 0.64	356	119	0.45	0.35 to 0.57
*Optimal chemotherapy***												
No	160	102	1	Reference	229	156	1	Reference	208	145	1	Reference
Yes	316	97	0.42	0.29 to 0.62	534	239	0.54	0.41 to 0.71	673	306	0.46	0.34 to 0.62
FIGO stage												
I–IIA	124	15	1	Reference	183	24	1	Reference	181	41	1	Reference
IIB–IV	352	184	6.64	3.64 to 12.10	580	371	7.64	4.71 to 12.40	700	410	3.21	2.24 to 4.61
Age (continuous)												
per 5 years	476	199	1.11	1.03 to 1.20	763	395	1.08	1.03 to 1.14	881	451	1.02	0.97 to 1.07
ECOG performance												
0/1	374	126	1	Reference	564	239	1	Reference	657	286	1	Reference
>1	102	73	1.71	1.11 to 2.63	199	156	1.64	1.25 to 2.16	224	165	1.59	1.25 to 2.02
Ascites												
≤500 ml	284	88	1	Reference	413	142	1	Reference	455	160	1	Reference
>500 ml	192	111	1.44	0.99 to 2.08	350	253	1.70	1.34 to 2.17	426	291	1.70	1.33 to 2.18
Comorbidity												
None	360	122	1	Reference	561	250	1	Reference	587	269	1	Reference
Present	116	77	1.39	1.01 to 1.81	202	145	1.61	1.23 to 2.11	294	182	1.26	1.01 to 1.57
Histology												
Other	147	47	1	Reference	282	138	1	Reference	265	109	1	Reference
Serous	329	152	1.23	0.83 to 1.81	481	257	0.83	0.64 to 1.07	616	342	0.92	0.72 to 1.19
Grade												
G1/G2	261	89	1	Reference	409	189	1	Reference	382	168	1	Reference
G3/G4	215	110	1.07	0.78 to 1.45	354	206	0.91	0.74 to 1.13	499	283	0.99	0.82 to 1.19

Note: *Complete surgery: No missing staging item (FIGO I–IIA)/0 mm tumor residual (FIGO IIB–IV)

**Optimal chemotherapy: Platinum-based (FIGO I–IIA, except for FIGO IA G1)/Platinum-taxane (FIGO IIB–IV)

3.5.4 Adherence to both surgical and chemotherapy guidelines

For the assessment of overall guideline adherence, both surgery and chemotherapy were combined into categorical variables with four categories (OP+/CT+, OP+/CT-, OP-/CT+, OP-/CT-), three categories (OP+/CT+, OP+/CT- or OP-/CT+, OP-/CT-) and two categories (OP+/CT+, non-standard). In the first sensitivity analysis, "optimal" surgery and standard chemotherapy were taken together to define the therapeutic standard for treatment of ovarian carcinomas. In a second sensitivity analysis, the stronger "complete" definition of surgery was combined with standard chemotherapy. The summary results will be presented separately for early-stage and advanced-stage disease. Then all stages of disease will be pooled and the predictors of overall guideline adherence will be determined. Finally, patient survival according to adherence to the combined treatment guidelines will be described in more detail and mediation models including one categorical mediator with two and three categories will be presented and compared to the primary mediation analysis with two binary mediators.

3.5.4.1 Early-stage ovarian cancer

In the primary analysis, the definition of adherence to surgical guidelines allowed for maximally one missing staging step ("optimal staging"). This definition was AND-combined with the chemotherapy standard according to which no chemotherapy is recommended for patients with FIGO IA G1 tumors, and adjuvant platinum-based chemotherapy is recommended for the rest of the patients with early ovarian cancer. In line with this definition, patients with early ovarian cancer had a higher chance of receiving appropriate treatment in the later cohorts: In 2001, both staging (i.e., maximally one staging step was omitted) and chemotherapy were optimal in only 19% of patients. This proportion increased to 33% in 2004 and to 55% in 2008. Simultaneously, the proportion of patients neither optimally staged nor treated in accordance with chemotherapy recommendations decreased over time (29% in 2001, 25% in 2004, and 14% in 2008). In all three cohorts, only a small number of patients received optimal surgery but were not treated appropriately with regard to chemotherapy (3% to 6%). Figure 26 shows the adherence to treatment guidelines combined for optimal surgery and chemotherapy in trial versus non-trial hospitals by year of diagnosis. In the first and the last cohort, the patients in trial hospitals were more likely to be optimally treated surgically and chemotherapeutically than patients in non-trial hospitals (30% vs. 10% in 2001, 58% vs. 50% in 2008). In contrast, in QS-OVAR 2004, 28% of patients in trial hospitals but 38% of patients in non-trial hospitals were treated according to the definition given above. However, in this second cohort, more patients in trial hospitals compared to non-trial hospitals received standard chemotherapy without being optimally staged (44% vs. 28%).

In a further sensitivity analysis, the stronger "complete" definition of surgical staging was applied; as outlined above, this definition did not allow for any missing staging step. Based on this stronger definition, only 6% of patients in OS-OVAR 2001 received both complete staging and standard chemotherapy. These numbers increased to 15% in 2004 and 31% in 2008. In all three cohorts, most of the patients were treated according to chemotherapy guidelines but were not completely staged (62% in 2001, 55% in 2004, and 49% in 2008). Again, only a few patients received complete staging with all nine steps performed but were not treated with the chemotherapy guidelines (4% in 2004 and 3% in 2008). As can be seen from Figure 27, in all three cohorts, patients in trial hospitals had a slightly higher chance to be treated according to both surgical and chemotherapeutical guidelines than patients in non-trial hospitals (9% vs. 3% in 2001, 17% vs. 13% in 2004, and 33% vs. 27% in 2008).

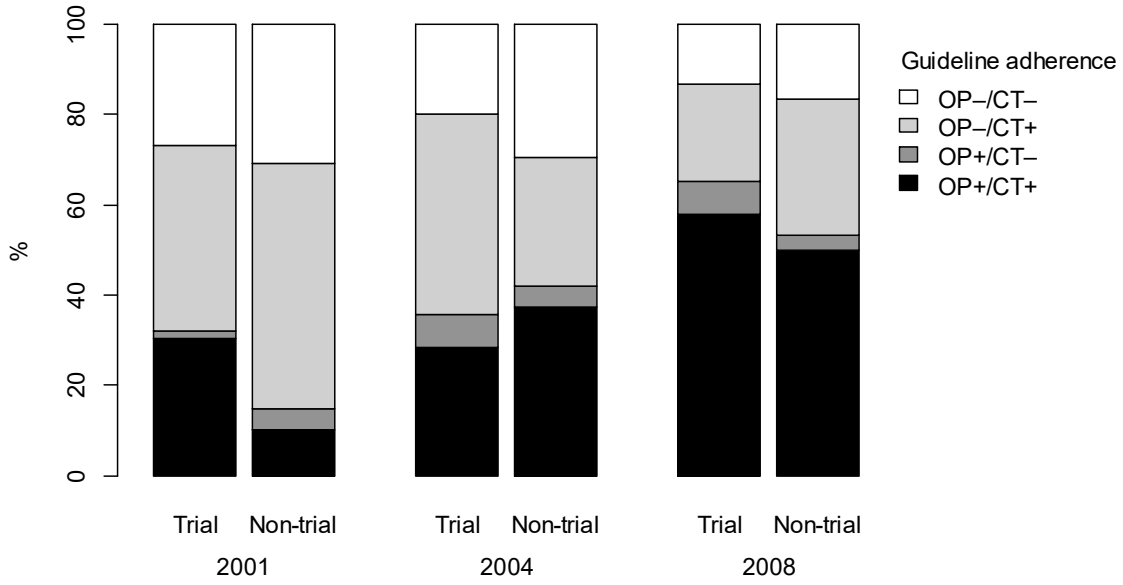

Figure 26: Adherence to surgical (optimal) and chemotherapy guidelines in early-stage ovarian cancer in trial and non-trial hospitals by year of diagnosis

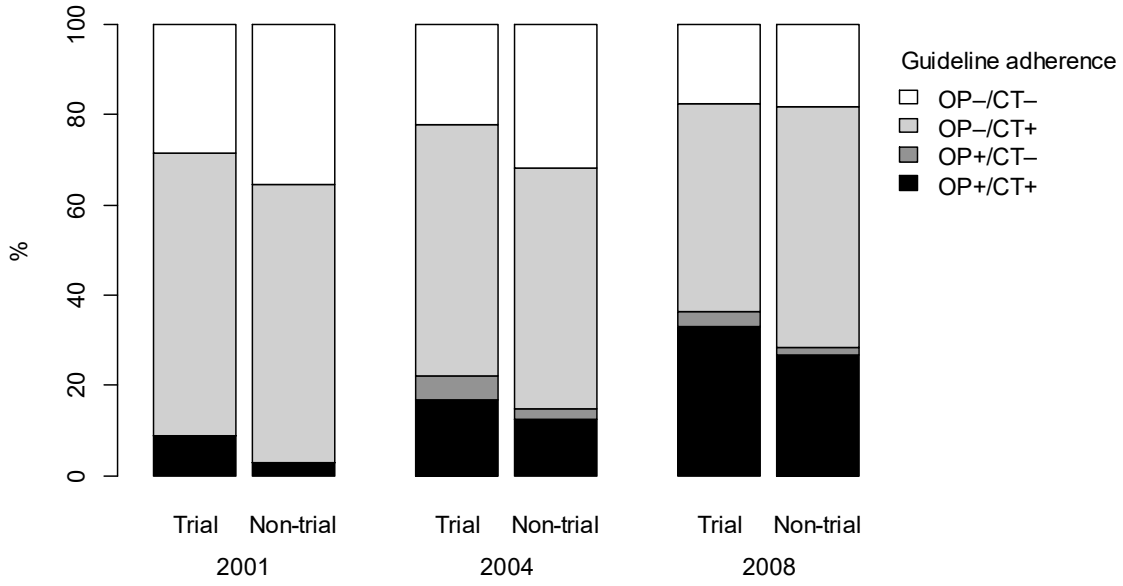

Figure 27: Adherence to surgical (complete) and chemotherapy guidelines in early-stage ovarian cancer in trial and non-trial hospitals by year of diagnosis

3.5.4.2 Advanced-stage ovarian cancer

In patients with advanced disease, guideline adherence was defined as "optimal" debulking with a maximum of 10 mm residual tumor, combined with platinum-taxane chemotherapy. According to this definition about 45% of patients in QS-OVAR 2001, 47% of patients in 2004 and 54% of patients in 2008 received both the optimal surgical and chemotherapeutic treatment. About 20% of patients in each cohort received the recommended chemotherapy but had tumor residuals after surgery of more than 10 mm. The proportion of optimally debulked patients without the appropriate chemotherapy treatment decreased over the years from 16% to 11%. There were about 18% of patients in 2001, 17% in 2004, and 13% in 2008 with neither recommended surgical nor chemotherapy treatment. In all three cohorts, patients in trial hospitals had a higher chance to be optimally debulked and to receive the recommended platinum-taxane chemotherapy (50% vs. 38% in 2001, 54% vs. 38% in 2004, and 60% vs. 45% in 2008, Figure 28).

According to the stronger definition, "complete" debulking was required together with standard platinum-taxane chemotherapy. This stricter treatment standard was fulfilled in only 24% of patients in 2001, in 29% of patients in 2004, and in 35% of patients in 2008, respectively. Here again, patients in trial hospitals had a higher chance to be treated in accordance to treatment guidelines compared to patients in non-trial hospitals (26% vs. 21% in 2001, 33% vs. 25% in 2004, 39% vs. 30% in 2008). About 40% of the patients in each cohort had tumor residuals after surgery but were at least treated with platinum-taxane chemotherapy. About one fourth (in 2001) to one fifth (in 2008) of the patients had tumor residuals after surgery and also did not receive the recommended chemotherapy. For more details see Figure 29.

3.5.4.3 Prognostic model for adherence to treatment guidelines

For the assessment of overall guideline adherence, both surgery and chemotherapy were combined, and a single binary outcome was used that reflected adherence to both treatment guidelines in the sense that both the chemotherapy and surgery (optimal or complete) recommendations were followed (Table 15 and 16).

In all three QS-OVAR cohorts, the patients who were treated in hospitals participating in clinical trials had a significantly higher chance to be treated according to the guidelines compared to those patients who were treated in non-trial hospitals. When optimal surgery was combined with standard chemotherapy, the adjusted odds ratios were 2.16 in QS-OVAR 2001 (95% CI 1.32 to 3.54), 1.66 in QS-OVAR 2004 (95% CI 1.15 to 2.38), and 1.47 in QS-OVAR 2008 (95% CI 1.05 to 2.05). Similar effects, however statistically not significant, were observed when complete surgery was combined with standard chemotherapy. Lower age, advanced stage, better performance status, and absence of ascites and comorbid conditions were associated with better adherence to surgical and chemotherapy guidelines. The other predictors showed rather inconsistent effects (see Table 15 and 16 for details).

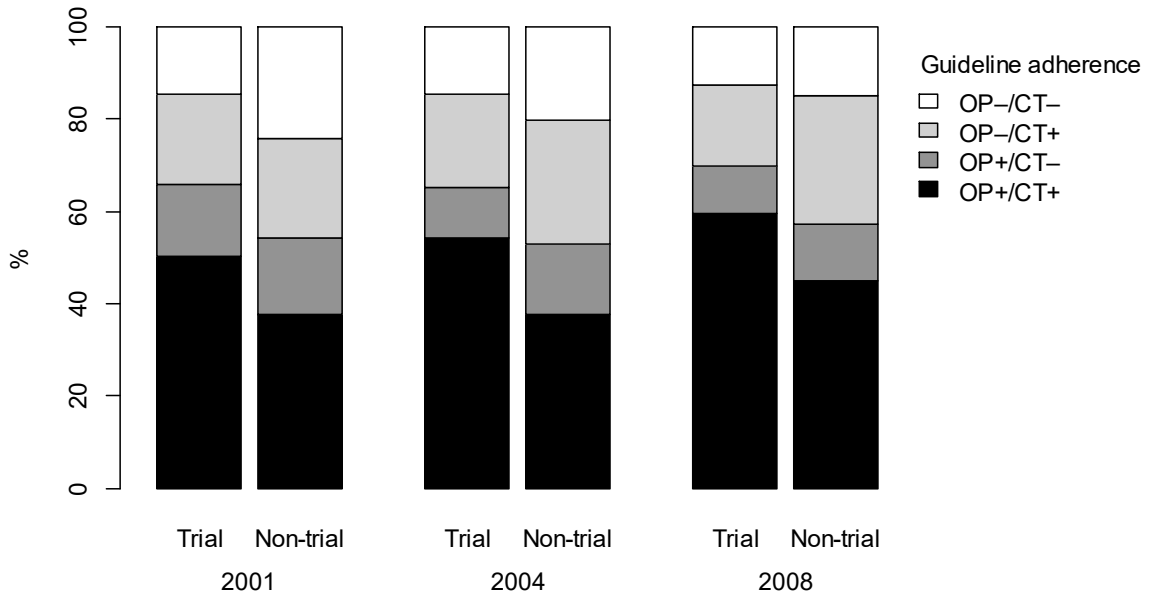

Figure 28: Adherence to surgical (complete) and chemotherapy guidelines in advanced-stage ovarian cancer in trial and non-trial hospitals by year of diagnosis

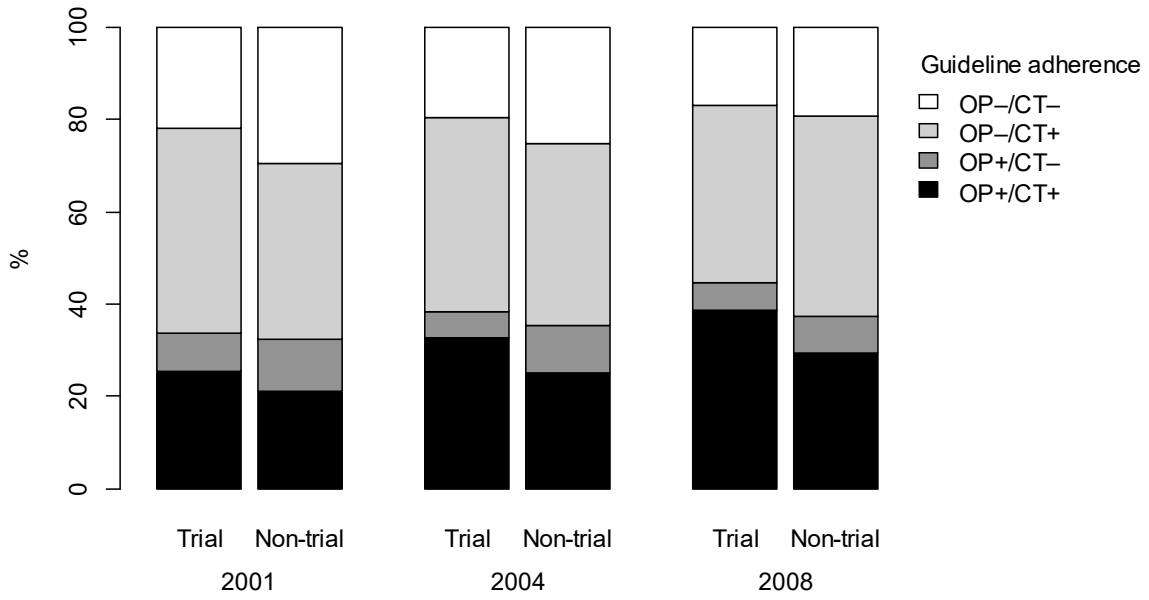

Figure 29: Adherence to surgical (complete) and chemotherapy guidelines in advanced-stage ovarian cancer in trial and non-trial hospitals by year of diagnosis

Table 15: Predictors of optimal surgery* + optimal chemotherapy**

	QS-OVAR 2001					QS-OVAR 2004					QS-OVAR 2008				
	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI
Research Activity															
Non-trial hospital	201	57	(28.4)	1	Reference	350	132	(37.7)	1	Reference	311	143	(46.0)	1	Reference
Trial hospital	275	127	(46.2)	2.16	1.32 to 3.54	413	200	(48.4)	1.66	1.15 to 2.38	570	337	(59.1)	1.47	1.05 to 2.05
FIGO stage															
I–IIA	124	24	(19.4)	1	Reference	183	60	(32.8)	1	Reference	181	100	(55.2)	1	Reference
IIB–IV	352	160	(45.5)	6.87	3.19 to 14.78	580	272	(46.9)	3.46	2.09 to 5.72	700	380	(54.3)	1.75	1.13 to 2.73
Age (continuous)															
per 5 years	476	184	(38.7)	0.79	0.72 to 0.87	763	332	(43.5)	0.80	0.74 to 0.87	881	480	54.5	0.75	0.69 to 0.81
ECOG performance															
0/1	374	174	(46.5)	1	Reference	564	291	(51.6)	1	Reference	657	422	(64.2)	1	Reference
>1	102	10	(9.8)	0.18	0.09 to 0.37	199	41	(20.6)	0.39	0.24 to 0.61	224	58	(25.9)	0.38	0.25 to 0.59
Ascites															
≤ 500 ml	284	110	(38.7)	1	Reference	413	193	(46.7)	1	Reference	455	289	(63.5)	1	Reference
> 500 ml	192	74	(38.5)	0.74	0.48 to 1.16	350	139	(39.7)	0.59	0.40 to 0.88	426	191	(44.8)	0.44	0.32 to 0.62
Comorbidity															
None	360	162	(45.0)	1	Reference	561	284	(50.6)	1	Reference	587	371	(63.2)	1	Reference
Present	116	22	(19.0)	0.50	0.27 to 0.93	202	48	(23.8)	0.43	0.28 to 0.65	294	109	(37.1)	0.55	0.39 to 0.78
Histology															
Other	147	51	(34.7)	1	Reference	282	116	(41.1)	1	Reference	265	140	(52.8)	1	Reference
Serous	329	133	(40.4)	0.98	0.57 to 1.67	481	216	(44.9)	0.90	0.64 to 1.26	616	340	(55.2)	1.25	0.87 to 1.80
Grade															
G1/G2	261	95	(36.4)	1	Reference	409	169	(41.3)	1	Reference	382	202	(52.9)	1	Reference
G3/G4	215	89	(41.4)	0.99	0.59 to 1.66	354	163	(46.0)	1.28	0.91 to 1.80	499	278	(55.7)	1.18	0.86 to 1.61

Note: *Optimal surgery: max. 1 missing staging item (FIGO I–IIA)/less than 10 mm tumor residual (FIGO IIB–IV)

**Optimal chemotherapy: Platinum-based (FIGO I–IIA, except for FIGO IA G1)/Platinum-taxane (FIGO IIB–IV)

Table 16: Predictors of complete surgery* + optimal chemotherapy**

	QS-OVAR 2001					QS-OVAR 2004					QS-OVAR 2008				
	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI	Total	Adhered	(%)	OR	95% CI
Research Activity															
Non-trial hospital	201	30	(14.9)	1	Reference	350	77	(22.0)	1	Reference	311	90	(28.9)	1	Reference
Trial hospital	275	61	(22.2)	1.32	0.79 to 2.19	413	120	(29.1)	1.48	0.96 to 2.28	570	214	(37.5)	1.25	0.87 to 1.79
FIGO stage															
I–IIA	124	7	(5.6)	1	Reference	183	27	(14.8)	1	Reference	181	56	(30.9)	1	Reference
IIB–IV	352	84	(23.9)	12.72	4.87 to 33.24	580	170	(29.3)	5.23	2.92 to 9.38	700	248	(35.4)	2.96	1.84 to 4.77
Age (continuous)															
per 5 years	476	91	(19.1)	0.80	0.72 to 0.89	763	197	(25.8)	0.86	0.79 to 0.93	881	304	(34.5)	0.79	0.73 to 0.85
ECOG performance															
0/1	374	84	(22.5)	1	Reference	564	173	(30.7)	1	Reference	657	267	(40.6)	1	Reference
> 1	102	7	(6.9)	0.43	0.17 to 1.09	199	24	(12.1)	0.45	0.25 to 0.79	224	37	(16.5)	0.56	0.33 to 0.93
Ascites															
≤ 500 ml	284	69	(24.3)	1	Reference	413	130	(31.5)	1	Reference	455	209	(45.9)	1	Reference
> 500 ml	192	22	(11.5)	0.24	0.14 to 0.43	350	67	(19.1)	0.36	0.22 to 0.58	426	95	(22.3)	0.27	0.19 to 0.38
Comorbidity															
None	360	81	(22.5)	1	Reference	561	165	(29.4)	1	Reference	587	235	(40.0)	1	Reference
Present	116	10	(8.6)	0.52	0.24 to 1.09	202	32	(15.8)	0.58	0.36 to 0.92	294	69	(23.5)	0.66	0.43 to 1.00
Histology															
Other	147	28	(19.0)	1	Reference	282	70	(24.8)	1	Reference	265	94	(35.5)	1	Reference
Serous	329	63	(19.1)	0.77	0.45 to 1.34	481	127	(26.4)	0.80	0.54 to 1.18	616	210	(34.1)	0.95	0.68 to 1.33
Grade															
G1/G2	261	43	(16.5)	1	Reference	409	99	(24.2)	1	Reference	382	127	(33.2)	1	Reference
G3/G4	215	48	(22.3)	1.37	0.77 to 2.44	354	98	(27.7)	1.12	0.77 to 1.64	499	177	(35.5)	1.15	0.83 to 1.59

Note: *Complete surgery: no missing staging item (FIGO I–IIA)/complete debulking (FIGO IIB–IV)

**Optimal chemotherapy: Platinum-based (FIGO I–IIA, except for FIGO IA G1)/Platinum-taxane (FIGO IIB–IV)

3.5.4.4 Survival according to adherence to treatment guidelines

Figure 30 and Figure 31 show the Kaplan-Meier survival curves by stage of disease and QS-OVAR cohort according to adherence to the combined surgery (optimal or complete) and chemotherapy guidelines. As can be seen, the pattern was quite consistent for patients with early and advanced ovarian cancer in all three cohorts, showing higher survival rates for patients treated according to guidelines.

3.5.5 Mediation analysis

The multiple pathways framework (Lange et al. 2014) assumes that the two mediators are fulfilled independently of each other, as well as that the two mediators operate separately of each other. Because it is generally difficult to test such assumptions within the same data set, sensitivity analyses were performed on the QS-OVAR 2001 data to assess the robustness of the results. In these sensitivity analyses only one mediator was used to represent adherence to treatment guidelines with regard to surgery and chemotherapy. In the first sensitivity analysis, a single binary mediator was used that reflected optimal adherence to treatment guidelines in the sense that both the chemotherapy and surgery were optimal. In a second sensitivity analysis, an ordinal mediator variable was created. This ordinal mediator was a single variable indicating whether none (neither chemotherapy nor surgery), one (optimal chemotherapy or optimal surgery), or both criteria for treatment adherence (optimal chemotherapy and optimal surgery) were met. As can be seen from Table 17, the effect estimates and confidence intervals from both sensitivity analyses are very close to the two-mediator solution of the primary analysis.

Table 17: Sensitivity mediation analysis (QS-OVAR 2001 only)

	Primary analysis		One binary		Three categories	
Effect	HR	95% CI	HR	95% CI	HR	95% CI
Trial hospital						
Total effect	0.57	0.42 to 0.77	0.57	0.42 to 0.78	0.57	0.42 to 0.78
Direct effect	0.66	0.48 to 0.91	0.64	0.46 to 0.89	0.66	0.48 to 0.90
Indirect effect	0.86	0.76 to 0.96	0.89	0.80 to 0.97	0.86	0.77 to 0.96
via Surgery	0.91	0.83 to 0.99				
via Chemotherapy	0.94	0.87 to 1.00				
Covariates						
Age (per 5 years)	1.22	1.13 to 1.32	1.22	1.13 to 1.33	1.22	1.13 to 1.33
Stage (IIB–IV)	5.04	2.79 to 9.09	4.81	2.72 to 10.1	4.76	2.68 to 10.1
ECOG (> 1)	2.08	1.44 to 2.99	2.07	1.42 to 3.27	2.07	1.43 to 3.26
Ascites (> 500 ml)	1.69	1.20 to 2.38	1.67	1.19 to 2.46	1.68	1.19 to 2.47
Comorbidity (present)	1.58	1.18 to 2.11	1.51	1.14 to 2.19	1.50	1.12 to 2.16
Histology (serous)	1.24	0.85 to 1.82	1.22	0.85 to 1.92	1.23	0.85 to 1.91
Grade (3/4)	1.13	0.83 to 1.54	1.14	0.80 to 1.57	1.14	0.80 to 1.58

In final sensitivity analyses, potential biases due to non-linear relationships, interactions between exposure, baseline variables and mediators, as well as possible misclassification of the mediators were investigated (code and results available online, see electronic supplementary material in Rochon et al. 2014).

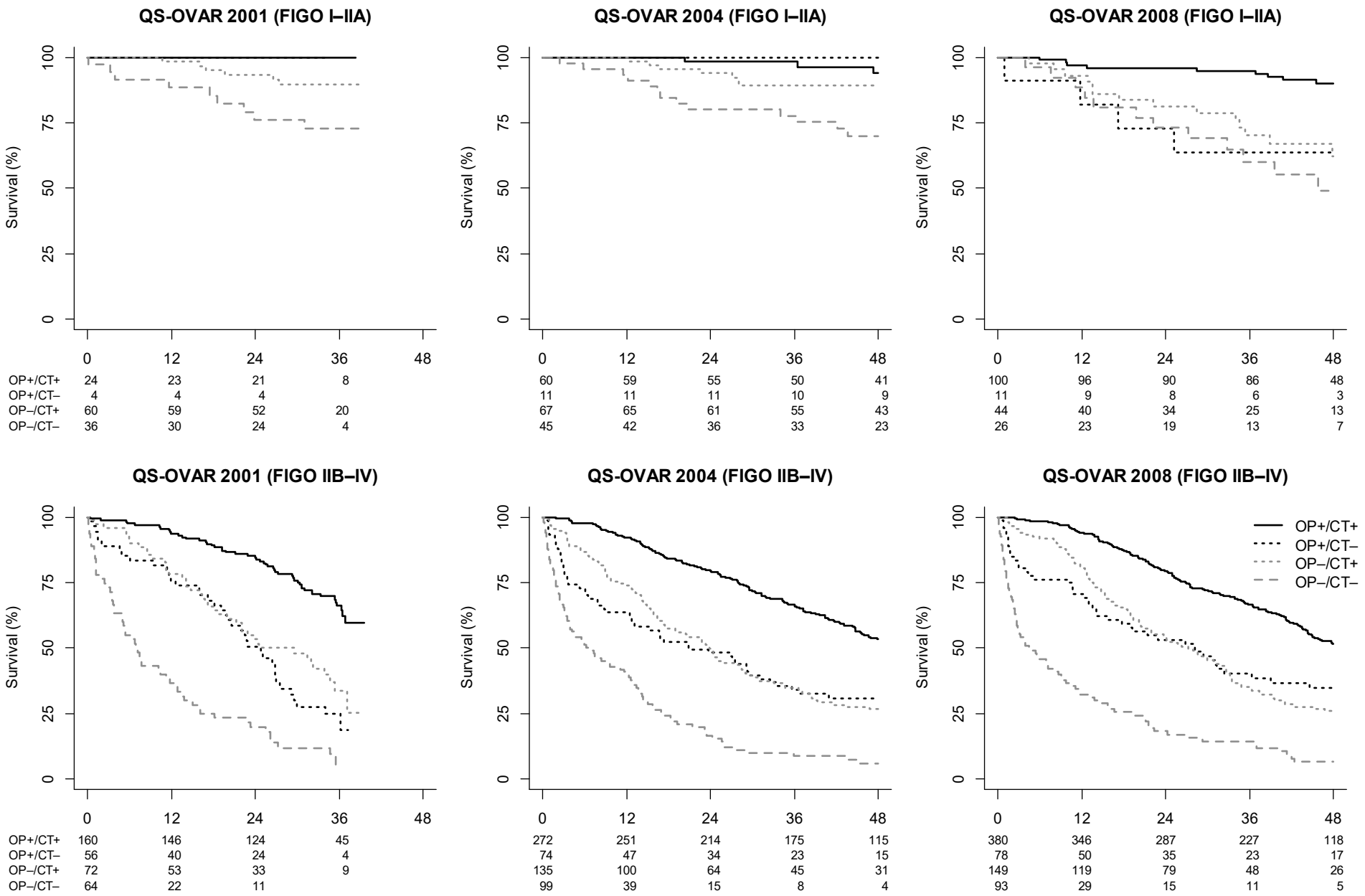

Figure 30: The relationship of patient survival and combined optimal surgery (OP) and optimal chemotherapy (CT) (for early-stage and late-stage ovarian cancer)

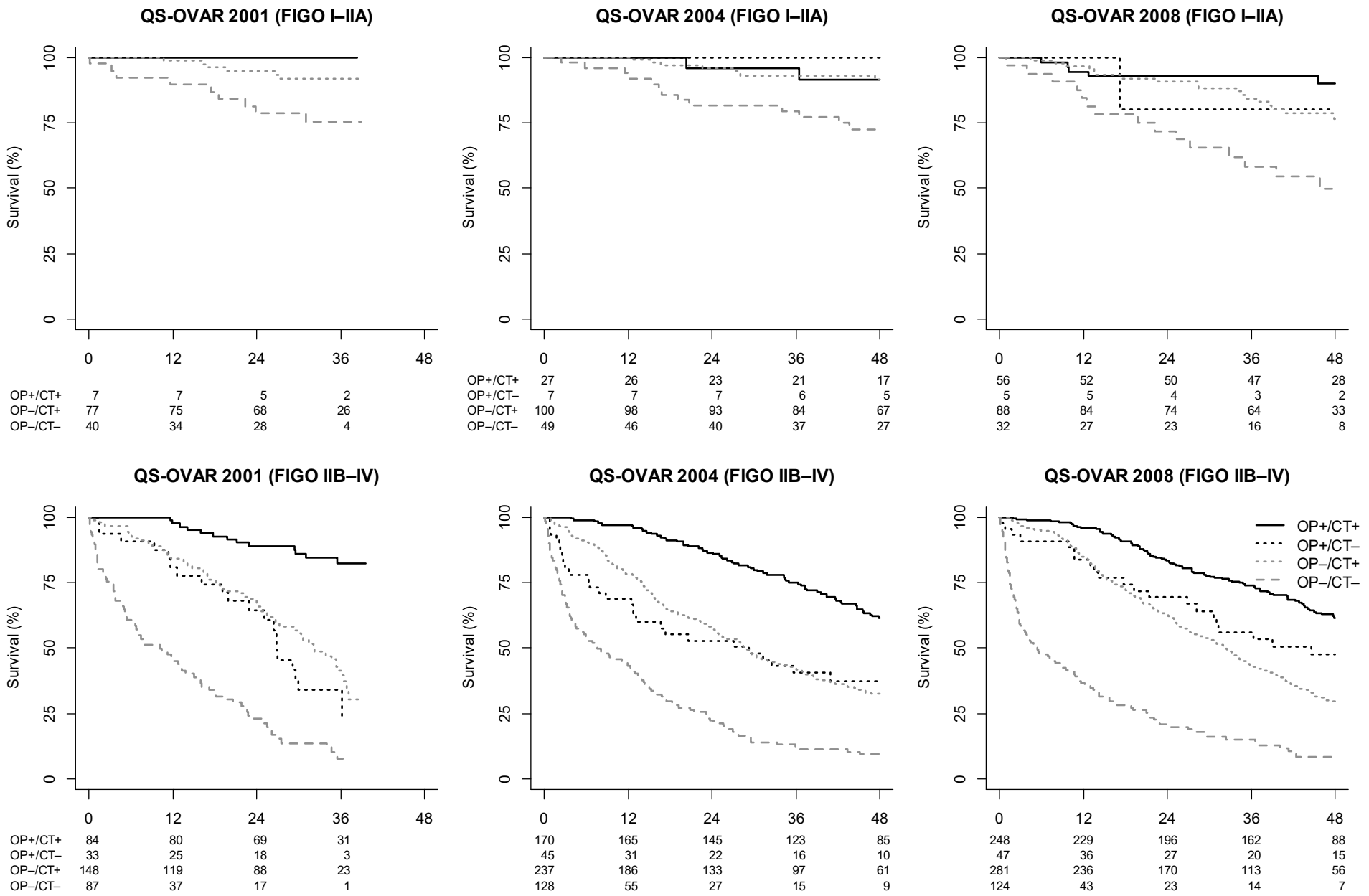

Figure 31: The relationship of patient survival and combined complete surgery (OP) and optimal chemotherapy (CT) (for early-stage and late-stage ovarian cancer)

4 Discussion

The present thesis integrates three key elements: a theoretical framework for assessing healthcare quality, a comprehensive collection and use of patient and hospital data from a nationwide quality assurance program, and a novel approach for evaluating causal mediation effects within a time-to-event setting. Each of these elements not only possesses its own merit, but all the facets complement each other, allowing a holistic view on the impact of institutional research activity on the quality of care and patient outcomes in ovarian cancer.

4.1 Effects of institutional research activity

Building upon the quality of health care assessment framework originally proposed by Donabedian (1966; 1980; 1988), this thesis investigated the hypothesis that that structure (hospital participation in clinical trials) influences the process of care (treatment), which in turn affects the outcome of care (patient survival). More specifically, it was anticipated that institutional clinical research activity would lead to better adherence to ovarian cancer treatment guidelines, subsequently improving care delivery and resulting in enhanced patient outcomes. Surgical treatment and chemotherapy administration were examined as potential mediators of the impact of hospital research activity on patient survival. The underlying mechanisms were explored using a new method for mediation analysis of time-to-event outcome data. The effects of institutional research activity were investigated using data from a German quality assurance program known as QS-OVAR. This program outlines care patterns in hospitals of different care levels and with varying patient volumes, also enabling comparisons between hospitals involved in clinical research within two German cooperative study groups and those not actively engaged in their study activities. The analysis utilized data from three cohorts of patients diagnosed with epithelial ovarian cancer in 2001, 2004, and 2008. Survival information was gathered for a minimum of three years after diagnosis in the first cohort and at least four years in the subsequent cohorts.

4.1.1 Survival

The data from the first QS-OVAR cohort supported the hypothesis that hospital participation in clinical research can improve the adherence to treatment guidelines of hospitals and can prolong patient survival. In QS-OVAR 2001, an independent relationship between hospital-level research activity and the survival outcome on patient level was observed (Table 9). Importantly, this positive association was not limited to patients actually enrolled in ongoing clinical trials but was extended to all patients treated in research-active institutions independent of their individual trial participation. The effect of hospital trial participation was most pronounced in patients with advanced stage of ovarian cancer (Figure 22, bottom left panel). Patients treated in trial hospitals were more likely to receive treatment according to guidelines compared to patients treated in non-trial hospitals (Tables 4–7). In addition, patients with “optimal surgery” and “optimal chemotherapy” lived longer than patients not treated according to the recommended standard of care (Table 8, Figures 20 and 21, see Section 2.1 for the definition of “optimal”, and discussion below).

Hence, it was logical to inquire how much of the observed benefit of hospital trial participation on patient survival was channeled through the appropriate use of surgical and chemotherapy treatments. To address this question, a new methodological approach for evaluating mediation in situations involving time-to-event outcome data was implemented and further refined (Lange et al. 2012; Lange et al. 2014; Rochon et al. 2014).

As shown in Table 10, the effect of research activity on patient survival was indeed partially mediated through better adherence to treatment guidelines in trial hospitals compared to non-trial hospitals. Taking into account several known patient and disease characteristics at baseline, the overall hazard ratio (total effect) of 0.57 was decomposed into a direct effect of hospital research activity of 0.66 and two indirect effects of 0.91 and 0.94 mediated through surgery and chemotherapy, respectively. The aggregate indirect effect through both mediators was 0.86, that is, about 27% (on the log hazard ratio scale) of the beneficial effect of research activity seemed to be mediated through both surgery and chemotherapy. In contrast, the "direct effect" of 0.66 summarizes a non-negligible bunch of mechanisms of unknown origin that operate in parallel to guideline adherence.

The beneficial effect of research activity, however, was only visible in QS-OVAR 2001 and was not replicated in the following cohorts (Tables 9 and 10, Figure 22). In QS-OVAR 2004 and 2008, the total effects of hospital-level research activity were 1.18 and 0.97, respectively. Both effects did not differ significantly from a hazard ratio of 1. In 2004, the still present but weak indirect effects (0.90, with 0.95 from surgery and 0.94 from chemotherapy) were masked by a much higher direct effect of unknown origin favoring research-inactive hospitals. In 2008, both direct and indirect effects were around 1, not pointing in any direction.

4.1.2 Adherence to treatment guidelines

In the first QS-OVAR cohort (Figure 7), less than 10% of patients received complete staging and less than 30% received "optimal" staging; the latter in the sense that maximally one staging procedure was omitted. In the same cohort, surgical results in advanced ovarian cancer (Figure 10) showed less than 40% of complete debulking and above 60% of "optimal" debulking (i.e., maximally 10 mm tumor tissue left). Despite these disappointing overall results, trial hospitals more frequently adhered to staging guidelines in patients with early ovarian cancer and also achieved complete or at least optimal debulking in more patients with advanced ovarian cancer (Tables 4 and 5, Figures 7 and 10). Similar, though weaker differences in favor of trial hospitals were observed in QS-OVAR 2004 and 2008 (Table 5).

In QS-OVAR 2001, adjuvant chemotherapy was administered to only about 65% of eligible patients (Table 6B/C). Although the difference was not large, trial hospitals more frequently adhered to treatment guidelines in advanced ovarian cancer and, to a higher extent, selected stage-appropriate chemotherapy regimens (Tables 6, 7). With time, however, both trial hospitals and non-trial hospitals improved slightly regarding adherence to chemotherapy guidelines and the originally observed beneficial effect of research-activity diminished.

Overall, the adherence to treatment guidelines improved gradually with each cohort of QS-OVAR (Tables 4 and 6). In particular, the surgical outcomes became better from 2001 to 2008. As expected, better adherence to treatment guidelines resulted in longer patient survival within each of the three cohorts (Table 8, Figures 20 and 21). The observed overall improvement in adherence to treatment guidelines did not, however, translate into even longer patient survival over the years (Figure 19).

4.2 Conceptual framework

Clinical research primarily provides the evidence needed to improve care and outcomes for future patients. Nevertheless, participation in clinical trials may also positively affect current patients. Past research on this topic mostly focused on individual patients and compared trial participants to non-trial participants. After a considerable amount of data had been gathered, however, it turned out to be difficult to establish a definite link between individual participation in clinical trials and outcomes (for reviews, see Braunholtz et al. 2001; Peppercorn et al. 2004; Vist et al. 2008). The consequence was a shift in emphasis from individual trial participation to organizational trial participation (Selby 2011), and the research question was reframed from "Do patients in clinical trials do better than patients outside trials?" to "Do health care institutions or service providers who are active in clinical research deliver better care and outcomes than those who are not?" (Selby and Autier 2011, p. 6). During the conceptual phase of this thesis, the latter question was comparatively new, and especially little attention had been paid to institutions active in clinical research. One single review on the impact of research engagement on the process and outcome of health care was available at that time (Clarke and Loudon 2011); it included 13 papers looking at the effects of health care practitioners' or institutions' participation in clinical trials. Despite promising signals for institutional research activity, however, the data was too limited to enable firm conclusions to be drawn and the evidence was less strong than originally thought to claim a positive relationship between research engagement and outcomes of care. The intriguing aspect of the new question was whether research-active hospitals deliver better outcomes for all their patients, that is, both trial participants as well as all the other patients treated in those hospitals. By definition, this question is of great importance for a much broader population of patients than the traditional approach (i.e., comparison of "on-trial" to "off-trial" patients). According to Pater et al. (2011), this question should be in the focus of attention of health care policy makers. Besides that, it is relevant to all patients who have to make decisions regarding the receiving of health care (i.e., where and by whom). The underlying assumption is that research activity might have a protocol-related impact on the care of patients who consent to enter clinical trials. In addition, engagement in clinical research might have a positive effect through the impact of research activities upon organizational infrastructure (e.g., staff, facilities). It is likely that this latter effect goes beyond the effect of direct individual involvement in clinical trials. From the policy makers' perspective, this aspect is also relevant because it relates to the justification of public resources spent for clinical research (e.g., Krzyzanowska et al. 2011). Finally, having a critical look at the question of research activity related to hospitals may reveal specific elements in the delivery of care that are actionable and accessible for improvement, independent of whether the institution is actually active in clinical research or not.

This thesis investigated the relationship between institutional research activity as defined by hospital participation in clinical trials, adherence to surgical as well as chemotherapy guidelines for ovarian cancer, and patient survival using Donabedian's (1966; 1980; 1988) theoretical framework of quality of health care as a starting point. Donabedian's framework consists of three factors: structure, process, and outcome. Structure refers to the context in which health care is delivered and can include preconditions such as buildings, equipment, staff including its qualifications and experiences as well as other characteristics such as accreditation status, patient volumes, hospital care level, research activity, teaching status or affiliation to a university. Process describes how this structure translates into practice. It sums up interactions between patients and health care providers and includes diagnoses, use of evidence-based treatments and preventive actions. Outcome refers to the effects of care and can encompass both hospital and individual results. As such, hospital standardized mortality ratios represent care outcomes at an institutional level, while direct patient outcomes like patient satisfaction,

alterations in patient health status, health-related quality of life, and patient survival signify care outcomes at an individual level. The greatest advantage of Donabedian's framework is that it is so general that it can be applied in various health care and delivery systems to questions of wide or narrow scope. At the basic level, the model can be used to assess and modify structures and processes within a health care delivery unit. At the higher level, the model can be applied to a large system to measure and to improve overall quality of care and outcomes for the entire patient population. However, potential limitations of Donabedian's approach have also been recognized (see Section 1.2 for a description of the three components including their advantages and disadvantages). Besides that, some authors (e.g., Mitchell et al. 1998) criticized the consecutive progress from structure to process to outcome as too simple and too linear for a framework. Consequently, the paradigm was judged as of limited utility for investigating how the three components interact in the nowadays complex and dynamic health care systems. A decade ago, Donabedian's framework was used to assess health care quality in ovarian cancer (Bristow et al. 2013). The authors concluded that due to the known limitations of Donabedian's approach, ideally, quality assessment and efforts of improvement should simultaneously incorporate structure, process and outcome, and balance them appropriately given the specific health care setting. Finally, it was acknowledged that the original framework does incorporate neither patient nor disease characteristics nor other factors that might be relevant for an adequate appreciation of the effects and underlying mechanisms within the process of health care delivery. Therefore, this thesis considered two extensions of Donabedian's quality of care triad.

Krzyzanowska et al. (2011) developed a framework specifically designed for a deep dive into how research activity might lead to better outcomes, even for those patients who are not directly involved in research projects. This framework emphasizes the interplay between infrastructure, care processes, and outcomes, while considering patient and disease characteristics. Despite its complexity, which may seem overwhelming for formal testing, it serves as a valuable guide for exploring potential mechanisms contributing to healthcare outcomes. A key aspect of this framework is the necessity for high-quality infrastructure, which includes not only resources such as staff, facilities, equipment, and financial support, but also the organization and coordination of these resources to conduct clinical trials. This aligns with the International Conference on Harmonisation (ICH) guideline on Good Clinical Practice (GCP), which since its inception, has stressed the importance of qualified staff in conducting clinical trials. The ICH GCP guideline has further emphasized this in its latest revision, highlighting the crucial role of competent personnel in ensuring the integrity and success of clinical trials. According to GCP, clinical investigators "should be qualified by education, training and experience to assume responsibility for the proper conduct of the trial and should provide evidence of such qualifications", they "should be familiar with the appropriate use of the investigational product(s)", and they "should have sufficient time, an adequate number of available and qualified staff, and adequate facilities for the foreseen duration of the trial to conduct the trial properly and safely" (ICH E6(R3) Guideline 2023, p. 11). Krzyzanowska et al. (2011) also noted that clinical research is seen as prestigious and promotes the recruitment and retention of high-quality staff. This includes not only doctors but also staff responsible for trial management, data handling, and quality assurance. This highly qualified staff is a key resource and a crucial element of infrastructure. As a result of the emphasis on qualified staff and adherence to GCP, it is reasonable to assume that all patients, including those not enrolled in trials, may benefit from the broader improvement of care that arises when research activity has a more general impact on care through the recruitment of more knowledgeable and experienced staff (Selby and Autier 2011). This improvement extends beyond personnel to encompass the overall infrastructure of an institution.

The sponsor of a clinical trial is responsible for ensuring that the trial is conducted in accordance with GCP, which involves selecting qualified trial personnel, providing a protocol, ensuring compliance with the trial protocol, and supplying facilities and appropriate equipment. GCP guidelines dictate that trial sites must have adequate facilities to support the trial and handle potential emergencies. This necessitates not only physical space for patient investigation but also proper equipment for conducting these investigations. To comply with GCP guidelines, many clinical trial sites must invest resources and adopt standard operating procedures that are followed when conducting a trial. Moreover, many trials require participants to undergo additional procedures, including new tests and tools according to the trial protocol. Research sponsors often fund these additional assessments, leading to the reasonable assumption that engaging in clinical research helps develop an institution's infrastructure to a higher level. While some elements of this infrastructure are applied directly to trial participants, access may extend to non-trial participants, especially after a specific clinical trial is completed. In addition to GCP recommendations, other initiatives, such as the 2008 statement from the American Society for Clinical Oncology on minimum standards and attributes of exemplary clinical trial research sites (Zon et al. 2008), provide guidance on clinical trial research sites. These structural requirements include formal maintenance of high educational standards, quality assurance, diversification of a trial portfolio and high accrual, clinical trial awareness programs, and financial oversight.

Linking clinical research conduct to outcomes requires considering variables related to processes of care. In Clarke and Loudon's (2011) review, practitioners and institutions participating in trials attributed infrastructure effects to better adherence to treatment guidelines and improved use of state-of-the-art evidence. Stiller (1994) highlighted that in oncology, both specialized centers (infrastructure) and standardized treatment (care processes) are crucial determinants of patient survival, often associated with participation in multicenter controlled trials. Later articles, in line with Krzyzanowska et al. (2011), emphasized the need for focusing on care quality, particularly care processes according to clinical practice guidelines and their impact on patient outcomes (Reade and Elit 2012; Reade and Elit 2019). The underlying assumption is that a process of care (e.g., appropriate chemotherapy)—when established by high-quality clinical trial evidence and applied to an individual patient—will eventually result in a good outcome for the patient. Support for an association between process and outcome usually comes from clinical trials in which a change in a certain process has led to an improvement in an outcome. The expectation is that when processes or interventions translate into routine care, outcomes will improve as well. Assessment of adherence to treatment guidelines and pathways is, thus, one of the first steps to consider in activities related to quality improvement. The process part assumes that practice patterns and health care delivery differ systematically between research-active and research-inactive institutions. For example, institutions engaged in research may be faster in incorporating new findings from the literature into clinical practice. With respect to technical aspects, research-active institutions may provide more accurate and timelier diagnoses or may be more likely to follow clinical treatment guidelines, adoption of which has been associated with improved outcomes. In specialized centers, for example, patients undergoing more complete staging procedures tend to be classified into higher cancer stages than in non-specialized centers. Participation in clinical trials may also influence physicians so that they transfer treatment modalities from trials into daily practice faster. It is also plausible that healthcare professionals in research-active environments provide superior care and outcomes due to skill enhancement, specialization, and personal attributes such as intrinsic motivation and medical engagement. These traits, combined with professional development in a research-intensive setting, may contribute to overall improvement in patient care and outcomes.

A structural variable often cited as a predictor of improved surgical outcomes is subspecialty training of the surgeon. For example, in the management of ovarian cancer, better process measures and patient outcomes were reported for patients managed by gynecologic oncologists compared to general surgeons (du Bois et al. 2009b). The most recent Ovarian Cancer Consensus Conference recommendations state explicitly that after initial diagnosis of advanced disease, patients should be assessed for primary surgery by a qualified gynecologic oncology surgeon (Karam et al. 2017; Vergote et al. 2022).

Survival is the most relevant patient related outcome measure in cancer care. Both patient and disease characteristics equally affect survival, and it is plausible to assume that patients differ systematically between the institutions in which care is delivered. When comparing research-active with research-inactive institutions in terms of patient survival, adjusting for potential confounders is thus essential. Social determinants of health such as education and income are associated with outcome in many areas of medicine including oncology (e.g., Krzyzanowska et al. 2011). Less educated or economically disadvantaged patients often experience worse outcomes: They present to the doctor at later stages, have poorer general health, limited access to care, inadequate health insurance, or they simply do not navigate the system as well as patients with higher socioeconomic status. Conversely, academic and research-active institutions have the reputation of providing better care. As a result, more educated, affluent patients or better "self-advocates" may be more likely to seek care in these institutions such that the patient mix may differ between research-active and research-inactive institutions. In contrast, patients with more severe diseases may also be more likely to be referred to research-active institutions, indicating that patient selection and referral biases can work both ways. In summary, adjusting for case mix, including both sociodemographic and disease characteristics, is crucial in any study evaluating the quality of care across institutions. Additionally, multivariable mediation analysis is mandatory to ascertain whether the impact of research activity is superficial or genuinely causal.

Expanding upon the core principles of Donabedian's triad of health care, Krzyzanowska's first extension highlighted the importance of various factors influencing health outcomes. In contrast, Selby et al. (2019a; b) introduced a second extension to the triad, specifically emphasizing the impact of research-active clinical teams on enhancing patient outcomes. In their model, a series of different mechanisms link research to better outcomes with the research-active clinical team being the central route by which outcomes can be improved. First, research-active teams generate new questions using evidence from literature. In addition, they apply the existing evidence as best practice guidelines and translate them into their daily work. According to the authors, staying informed of state-of-the-art developments should improve the process of care and, in turn, the patient outcomes. Exposure to best practices during the conduct of clinical research should lead to additional benefits. Finally, translational research results in further innovations that are then eventually tested in the experimental arms of clinical trials. When such trials are positive, they generate new evidence for best practice, which in itself will improve patient outcomes—both for those patients enrolled in trials as well as for the much larger number of patients whose treatment is improved through the process of the uptake of innovations. Selby et al.'s model assumes that this occurs more rapidly in research-active institutions. Finally, research-active clinical teams contribute to the development and refinement of medical practices by staying up-to-date with the latest advancements, which in turn promotes a culture of continuous learning and improvement within health care organizations, ultimately benefiting patients and the medical community as a whole.

All of the above-described conceptual frameworks will be used in the following paragraphs to explain some of the results of this thesis.

4.3 Explaining the results

In QS-OVAR 2001, it has been claimed that ovarian cancer patients treated in trial hospitals are more likely to receive treatment according to guidelines and have therefore better survival outcomes than patients treated in non-trial hospitals (du Bois et al. 2005a; Rochon and du Bois 2011; Rochon et al. 2014). These results, however, were not replicated in the following cohorts QS-OVAR 2004 and 2008.

First, let's assume there was a real effect of research activity in QS-OVAR 2001. The question then arises as to why hospitals that participated in clinical trials of the two German study groups were able to deliver better health care outcomes for their ovarian cancer patients compared to hospitals that were not involved in clinical research at that time. Some findings from the first cohort support the observed survival advantage in trial hospitals. Specifically, the consistently better surgical and chemotherapy treatments in trial hospitals as opposed to non-trial hospitals hold for patients with early as well as advanced ovarian cancer. They are in line with the assumption that research activity contributed to enhanced treatment, and this translated into better patient outcomes. A distinctive aspect of this thesis is the integration of the three components of Donabedian's model for quality of health care (structure, process, and outcome) within a single causal mediation analysis for time-to-event outcome data. The results of this mediation analysis revealed that approximately 27% of the beneficial effect of research activity on survival were mediated through surgery and chemotherapy.

Both estimated indirect effects are in line with the literature. For more than 20 years, it is well known that newly diagnosed ovarian malignancies profit most from a thorough surgical treatment. A publication at the time of analysis of the first cohort showed a 29% improvement in survival of patients with advanced ovarian cancer (HR = 0.71) that was triggered to a great extent by improvement in the proportion of optimal debulking from 43% to 66% (Bristow et al. 2002). Similarly, since the publication by Piccart et al. (2000), there is strong and confirmatory evidence for platinum-taxane combinations as standard treatment for patients with advanced ovarian cancer. When patient and disease characteristics including residual tumor were taken into account, addition of paclitaxel to a platinum chemotherapy resulted in a 27% reduction in the death rate (HR = 0.73). Bristow et al.'s and Piccart et al.'s findings were quite recent at the time of QS-OVAR 2001; and it is a plausible assumption that hospitals participating in clinical research took up this new evidence quicker than their non-trial counterparts (e.g., due to membership in the collaborative German study group network and the access to information from the international GCIG network). At that point in time, the participation in clinical trials offered the only possibility of nationwide multicentric quality assurance for an oncologic treatment.

In addition to the effects mediated by surgery and chemotherapy, patients may have benefited from general improvement of care even if they were not actually enrolled in trials because they were treated at trial hospitals. Based on the frameworks described above, it can be speculated that in 2001, the members of the two cooperative study groups that organize almost all national and international clinical trials for ovarian cancer in Germany achieved better outcomes for all their patients because of their higher standard compared to their peers outside of this collaborative network. This may arise if research activity has an impact on delivery of care by availability of highly qualified personnel, adequate infrastructure and best practices reflecting state-of-the art treatment guidelines. Research participation might then stimulate teams to improve the health care delivery by, for instance, quicker and better-informed decision making even for patients outside of the ongoing clinical trials. As clinical research continues to evolve and incorporate cutting-edge technologies, it is anticipated that the quality of care will experience further improvements. The more recent publications published by Boaz et al. (2015), Downing et al. (2017), Harding et al. (2017) and Selby et al. (2019a; b) support this view. Therefore, these publications will be described in more detail in the following paragraph, providing further insights into their findings and implications.

Boaz et al. (2015) conducted a so-called “hourglass review” with altogether 33 papers and concluded that it is reasonable to assume that when clinicians and health care organizations engage in research, there is a higher likelihood of improvement in their health care performance, even if this was not the primary objective of conducting research. In addition, the authors underlined that in the past, the measures used to assess the performance of health care organizations were primarily costs and rather basic performance measures (e.g., number of patients treated per year) while research activity was not considered. In the meantime, however, there has been an increasing interest in the latter characteristic, especially in the United Kingdom, and the role of research engagement was explored from various perspectives. For example, in 2008 the National Institute for Health Research began funding a major 5-year pilot research program of translational research in England, establishing nine “Collaborations for Leadership in Applied Health Research and Care” (CLAHRC). This program was motivated by the observation that previously, university researchers carried out studies with little involvement of staff and/or health service users. With time, the government and the public have sought to close the gap between the generation and exploitation of evidence, by promoting innovative culture, enabling partnerships and facilitating the uptake of novel evidence into public health practice. The expectation was that the closer the collaborations are the more successful the knowledge translation will be. A number of evaluations were carried out to examine whether the various collaborations worked as intended and if so, why (Heaton et al. 2015). At the end of the successful pilot, 13 new CLAHRCs were funded for another 5 years. The interesting aspect of this initiative is the focus on collaborative leadership and networks, which matches well the key characteristics of the German study group approach used in this thesis to define research activity at hospital level. A similar structural health care characteristic was discussed by Bristow et al. (2015a). Using data from a large retrospective population-based study the authors reported that National Cancer Institute Comprehensive Cancer Center status is an independent predictor of adherence to ovarian cancer treatment guidelines and improved ovarian cancer survival. Downing et al. (2017) reported results from a large population-based study and were able to show that being treated in a research-active hospital is strongly associated with better outcomes for patients with colorectal cancer. The size of the observed impact of research participation on survival was comparable to the entire patient population impact seen following a successful interventional trial and this effect was present even after adjustment for various factors which may be expected to affect the performance of different institutions such as case-mix and hospital case volume. At around the same time, Harding et al. (2017) investigated the relationship between “research culture” (i.e., research participation at hospital level) and organizational performance. The latter included various indicators such as patient mortality rates, levels of patient satisfaction, staff turnover, staff satisfaction and organizational efficiency. All of the reviewed eight studies reported a positive association between research activity and organizational performance. The authors concluded that a stronger research culture is associated with benefits to patients, staff and the organization. Finally, Selby et al. (2019a) called to action and asked clinical research to become “everyone’s business” (2019a, p. 5). In essence, this means that clinical research should become a common endeavor, with involvement and investment from all parties in the healthcare system, including researchers, professionals, patients, organizations, policy makers, and the public. The idea is that by making clinical research a shared concern, it can lead to more generalizable clinical trials, as the trials would be more representative of diverse patient populations and healthcare settings. This, in turn, could improve healthcare outcomes, as the findings of the research would be more applicable to a wider range of patients and situations.

Against this background, an open question is why the effects of QS-OVAR 2001 were not replicated, neither in QS-OVAR 2004 nor in QS-OVAR 2008.

The likelihood of surviving ovarian cancer primarily depends on patient characteristics, tumor biology, and treatment (Gupta et al. 2010). While the first two factors are unchangeable, treatment can be influenced and thus plays a critical role in efforts to improve ovarian cancer outcomes. Surgery and chemotherapy, informed by global guidelines, are essential in addressing all disease stages. These guidelines typically describe detailed surgical procedures and include stage-appropriate chemotherapy recommendations (e.g., Wagner et al. 2013; Staebler and Mayr 2017; Karam et al. 2017; Armstrong et al. 2019; Vergote et al. 2022). However, adherence to these treatment recommendations has been reported to be suboptimal, indicating opportunities for improvement (du Bois et al. 2009a; Bristow et al. 2012). Timely incorporation of these guidelines into daily practice, strengthened communication between healthcare providers, and enhanced patient education, can potentially increase adherence to these guidelines, ultimately leading to better patient outcomes.

Cancer staging is the process of finding how far the tumor has grown and spread into other parts of the body. Staging is performed surgically; it includes the description of the extent of the cancer and serves as the initial treatment for ovarian cancer. The goal of staging is thus twofold: to see how far the cancer has spread, and to remove all visible tumor tissue. Staging can therefore be considered both a diagnostic tool and a therapeutic procedure. Accurate surgical staging is critical to ensure appropriate subsequent treatment, as without staging women may suffer from uncontrolled proliferation of the disease. In early ovarian cancer, staging has already been reported in the 1980s to be an area with a wide translational gap, that is, standard procedures had not sufficiently been transferred into clinical routine (McGowan et al. 1985). In advanced ovarian cancer, primary debulking surgery followed by chemotherapy has become the standard of care since more than 40 years (Griffiths and Fuller 1978). The amount of residual disease after primary surgery is still the most powerful determinant for patient outcome with survival being best in patients with no residual tumor (Karam et al. 2017).

Treatment guidelines recommend postoperative chemotherapy for all patients with ovarian cancer, with the exception of FIGO stage IA grade 1. Patients with early-stage disease should receive platinum-based treatment, whereas patients with advanced-stage disease should receive a platinum-taxane combination (e.g., Karam et al. 2017; Vergote et al. 2022). These recommendations have not changed much since Ozols et al.'s (2003) publication reporting the results of a non-inferiority trial with 840 patients and establishing carboplatin as a replacement for cisplatin in the initial treatment of advanced ovarian cancer following primary debulking surgery. The regimen consisting of carboplatin and paclitaxel still represents the backbone of treatment in advanced ovarian cancer (Boyd and Muggia 2018).

Even though clinical recommendations for diagnosis, treatment and follow-up of ovarian cancer were available for both early and late ovarian cancer in early 2000s (Bauknecht et al. 2000; European Society for Medical Oncology (ESMO) Guidelines Task Force 2001), the adherence to treatment guidelines was quite low in the first QS-OVAR cohort compared to the adherence in the later cohorts. It seems to take some time for a new treatment guideline to be translated into practice. It might be speculated that hospitals participating in clinical trials were nevertheless faster in the uptake of these recommendations than non-trial hospitals and therefore adhered better to treatment guidelines in 2001 already. In the following years, however, it might be assumed that also the non-trial hospitals became more and more aware of the state-of-the art treatment and adopted the chemotherapy recommendations as well. Even though the observed adherence to treatment guidelines still offered room for improvement, especially after the study published by Ozols at al. (2003), carboplatin-paclitaxel was at least well known as standard chemotherapy treatment in advanced disease and the ovarian cancer community should have been aware of it.

One should consider, however, that there also might be other reasons for not being treated according to the standard recommended in guidelines. Adherence to treatment guidelines might be influenced by infrastructure characteristics (e.g., specialization of the physician, knowledge, skills, experience; hospital research activity level or volume) but can also be impacted by the availability of treatment according to local policies and reimbursement strategies. In addition, patients may decide not to undergo chemotherapy to avoid the heavy side effects of the treatment. Finally, disease characteristics can play a role in decisions on a treatment (e.g., age, comorbidities, poor prognosis). It is thus worth noting that the quality of staging in early-stage disease and tumor reduction in patients with advanced ovarian cancer improved markedly from the first to the last QS-OVAR cohort for both, trial and non-trial hospitals (Figures 7 and 10, see also Table 5). Though less pronounced, the same trend was observed for chemotherapy. The reasons for these overall improvements may be a mixture of all of the above. In addition, they might even be result of participation in the quality assurance program itself.

A further explanation for the discrepancy of the findings across cohorts is that the effects observed in the first cohort were simply overestimated. This phenomenon, where an effect is observed once and then never again, is not uncommon, particularly in pilot studies. These initial studies are often used to evaluate the preliminary effects of a treatment, which can lead to a misestimation of the true effects (Kistin and Silver-stein 2015). Specifically, small pilot studies are more prone to either overestimating or underestimating the actual effect. If the true effect size is underestimated, a planned subsequent larger study may be considered not worth the conduct and the potentially important true effect would remain undiscovered. If the true effect size is overestimated but still used to inform further research, resources are invested in subsequent studies, but these studies may be underpowered. In addition, overestimations can result in inappropriate conclusions and misguided decisions with serious implications for future research. Indeed, the effect observed in QS-OVAR 2001 was larger than effect sizes for overall survival in successful clinical trials of new effective chemotherapy treatments in ovarian cancer (e.g., HR = 0.73 reported by Piccart et al. 2000 or 0.75 reported by Katsumata et al. 2009). However, it must be taken into account that in clinical trials comparing chemotherapy treatments, multivariable models for overall survival always include residual disease as a baseline covariate that serves as an important prognostic factor for survival in these studies (Karam et al. 2017). In contrast, residual disease is a process variable in a quality of care model, and in this thesis, it was assumed that this process variable is also influenced by hospital-level research activity. Therefore, the survival hazard ratio for trial hospitals compared to non-trial hospitals was adjusted for all relevant patient and disease characteristics in multivariable analyses except for the surgical outcome (i.e., residual disease). The estimated effect in QS-OVAR 2001 might therefore be larger than effects reported in chemotherapy trials. Second, even though the effect of hospital-level research activity on patient survival might have been overestimated in the first QS-OVAR cohort, QS-OVAR 2001 was not the first cohort indicating benefits from hospital research activity. The first pilot evaluated the records of 501 patients treated between July and September 2000 in 87 institutions (du Bois et al. 2001a; b). Although this pilot did not assess patient survival, it outlined the patterns of care in ovarian cancer in Germany and identified areas with room for improvement in terms of adherence to treatment guidelines. The German data were found to be consistent with international results. Furthermore, the pilot highlighted differences between hospitals. Notably, patients treated in trial hospitals had a higher likelihood of receiving optimal treatment compared to those treated in non-trial institutions. These findings inspired the hypothesis that hospital-level research activity enhances adherence to treatment guidelines and, via this route, improves patient outcomes.

Even though the argument of a chance finding cannot be fully ruled out, there are other explanations that might help understand why the effects observed in QS-OVAR 2001 were not replicated in the following cohorts. Comparing results between hospitals can become challenging even with modern methodology because of the many known and unknown mechanisms that can affect outcomes and contribute to a trial effect. Braunholtz et al. (2001) differentiated already between treatment effects, protocol effects, care effects, Hawthorne effects, and placebo effects. The treatment effect, for example, would benefit patients if the new treatments tested in trials tended to be better than the standard treatment. Consequently, patients who received superior treatment would be shown to have directly benefited from participation in the trial. In contrast, the protocol effect would rather arise from the fact that trials follow protocols in which treatments are usually carefully described. If trial protocols improve delivery of care and thus advance the subsequent outcomes then one would probably observe a trial effect. The care effect would be difficult to estimate separately from the protocol effect but could result from additional trial-related care aspects.

Benefits for patients enrolled in study protocols in oncology have been reported earlier but several pitfalls interpreting these effects have been identified. The simplest explanation for the observed beneficial effect of hospital-level research activity in the first QS-OVAR cohort might be that patients enrolled in ongoing trials during the study period contributed to this effect, indicating a protocol effect. In that case, the effect should vanish when excluding trial patients in trial hospitals and comparing non-trial patients in trial hospitals to similar patients in non-trial hospitals. In the three QS-OVAR cohorts taken together, almost 60% of patients were treated in hospitals participating in clinical trials. However, only 14% (175 out of 1258) of these patients were actually enrolled in prospective clinical trials of the two German study groups. In the third quarter of 2001, there were 22% (59 of 275) of patients enrolled in ongoing trials in trial hospitals (versus, of course, 0% in non-trial hospitals). In QS-OVAR 2004, this percentage was even lower (6%), and in QS-OVAR 2008 about 16% of patients in trial hospitals actively participated in clinical trials. Even though it is tempting to conclude that the 22% of patients in QS-OVAR 2001 triggered the observed effect, it is unlikely that only those patients who actually participated in clinical trials caused an effect of the observed size. What speaks against the assumption that the observed effect in 2001 only reflects a kind of protocol effect? First, there were no trial protocols available for patients with early ovarian cancer at that point of time. Second, the only active German protocol for advanced ovarian cancer in 2001 was AGO-OVAR-7 (Pfisterer et al. 2006) did not show superiority for the experimental arm. Finally, in QS-OVAR 2001, almost 50% (80 of the 165) hospitals participated in cooperative study group activities. Even though the protocol effect as described above cannot be fully excluded, it is unlikely that the same physicians in the same hospitals delivered care differently to patients in the only active trial at that time and outside of this clinical trial, expect perhaps for a longer and more granular follow-up for trial patients. In addition, if the observed effect in favor of research activity in QS-OVAR 2001 would only reflect a care effect, a more pronounced positive effect of hospital trial participation should have also been observed at least in the last of the three QS-OVAR cohorts. In this last cohort, 16% of patients in trial hospitals were trial participants but no relevant direct or indirect trial effects were observed.

A further explanation might be the so-called Hawthorne effect (Franke and Kaul 1978). This effect means that behavior tends to change when individuals are under observation. The involvement in a clinical trial might affect both patients and clinicians (McCarney et al. 2007). A Hawthorne effect might also have occurred in QS-OVAR. The simple fact that hospitals were asked to participate in the quality assurance program and agreed to respond to the survey and document the required quality data, might have affected at least the participating hospital personnel who perhaps tried to show their best.

In fact, this mechanism can be seen as a welcome and beneficial byproduct of quality assurance policies (e.g., Manzanera et al. 2018; Breyer et al. 2019). A counter argument against this explanation is, however, that the hospitals were asked to provide their patient data approximately one year after the initial diagnosis. This is the time point when surgical treatment as well as primary chemotherapy are usually completed and patients are in the follow-up phase. As overall survival is an objective patient outcome, it is almost impossible that it might have been influenced just by the knowledge of being part of a quality assurance program. Nevertheless, the better results in adherence to treatment guidelines observed with time might be one of the major accomplishments of the program itself.

Perceived benefits from trial participation may also be due to potential biases, such as selection bias. This would mean that the advantage of participating in a trial could be just a result of the selection process, rather than a true reflection of the benefit for the broader population. One possible explanation for the observed results would be then that research-active institutions are systematically different from non-research-active hospitals. Hospitals participating in research might be self-selected, clinician-selected, and patient-selected. In 2001, for example, 50% of participating hospitals reported to be members of the two German cooperative study groups while only 25% of all German hospitals fulfilled this criterion. This could reflect a sampling bias. However, this did not change over time, and in the later cohorts, trial hospitals were equally overrepresented. One could argue that the hospitals participating in the quality assurance program were not representative for all German hospitals at all, as they are more willing to share their results—because they think they are better hospitals anyway. In this case, one could argue that the observed treatment patterns and guideline adherence in Germany would even be too optimistic and would be worse if all the German hospitals had been investigated.

Regarding other potential biases, there were some associations between hospital characteristics. Hospital trial participation was positively associated with higher hospital volume and higher hospital care level in all three cohorts of QS-OVAR investigated in this thesis (Figures 5 and 6). Reade and Elit (2012; 2019) reported the highest outcomes for high-volume providers in high-volume cancer centers, for example. One explanation for the relationship between hospital research activity and better outcomes could be the so-called "practice-makes-perfect" hypothesis. This hypothesis is based on the general belief that increased experience results in better skills, which in turn leads to better outcomes. In other words, the more a hospital or healthcare provider performs a specific procedure or treats a particular condition (i.e., the higher their volume), the better they become at it. This improvement can manifest as reduced complication rates, shorter hospital stays, lower mortality rates, and overall better patient outcomes. This hypothesis is often used to explain why high-volume hospitals and providers sometimes have better outcomes than their lower-volume counterparts. On the other hand, the observed relationship may also reflect that some hospitals are known as the best in their area of expertise, especially hospitals associated to universities; in consequence, they attract more patients or patients with specific disease characteristics. In this case, a referral selection bias could be a reason for the observed results. However, the multivariable analyses conducted in this thesis controlled for at least known potential confounders on patient level and adjusted for the other hospital characteristics. In addition, especially in 2004 and 2008, there were some non-trial hospitals with high patient volumes as well. These hospitals might have also gained experience simply by treating more patients per year. In this thesis, however, there was neither good evidence for better survival at high-volume hospitals nor at university hospitals (Tables 11 and 12, Figures 23 and 24). This is in line with the quite heterogeneous and weak evidence shown in a comprehensive review by du Bois et al. (2009b) for these hospital structure characteristics in ovarian cancer.

There might be another explanation possible for why the effect observed in the first cohort was not visible in the following two cohorts of QS-OVAR. There was a remarkable increase in the number of included patients and institutions from the first to the following two cohorts: QS-OVAR 2001 consisted of 476 patients from 165 hospitals, QS-OVAR 2004 included 763 patients from 245 hospitals, and QS-OVAR 2008 contained data on 881 patients from 240 hospitals. The quality assurance program was not designed as a longitudinal study, neither on patient level nor on hospital level. When looking at the participation of the individual hospitals in QS-OVAR over time, there were some overlaps between the cohorts: 67 hospitals entered all three cohorts while 123 hospitals participated twice. In contrast, 203 hospitals participated only once in QS-OVAR. With regard to the latter, 24% of the 203 hospitals were included only in the first cohort, 36% were included only in the second cohort and 40% entered only the third cohort. The definition of research activity was based on the participation in clinical trials of the Germany study groups in the respective study period. Because of this quite narrow definition, it cannot be entirely excluded that in the later QS-OVAR cohorts a few previously research-active hospitals were labelled as non-trial hospitals just because they did not participate in clinical trials according to the defined time period. In addition, as only trial activities covered by the two German study groups were considered for the definition of research activity, it might be argued that also some of the non-trial hospitals in 2004 and 2008 might have gained experience through other research activities. If this was true and non-trial hospitals participated in another kind of research, this could also result in improved adherence to treatment guidelines and better patient outcomes in these nontribal hospitals. The latter is, however, quite unlikely as the two study groups are both members of ENGOT and organize almost all clinical research activities in ovarian cancer in Germany. In addition, the effect of participation in trials of the two groups was deemed to cover more than just recruitment of patients in the third quarter of 2001, 2004 and 2008. To be considered as research-active, hospitals had to enroll patients in clinical trials before or during the third quarter of the respective year of diagnosis. Furthermore, it is important to note that the definition of research activity could be broadened to include other forms of research, not just clinical trials. This could potentially reveal a different landscape of research-active hospitals. Lastly, the impact of these other research activities on patient outcomes and guideline adherence could be a valuable area for future investigation. It is essential to consider the evolving landscape of research and the potential influence of various research activities on patient care. By exploring these factors, one could better understand the complex interplay between research, guideline adherence, and patient outcomes, ultimately leading to improved healthcare practices.

Correlation is not causation. Research-active hospitals may deliver better outcomes, but this could be due to superior performance and staff, not necessarily their research activity. It is therefore crucial to consider whether research activity truly improves infrastructure, leading to better performance and outcomes. Krzyzanowska et al. (2011) and Selby et al. (2019a; b) suggested that research activity can boost outcomes through various mechanisms. However, as systematic review concluded “when clinicians and health-care organisations engage in research there is the likelihood of a positive impact on health-care performance, but this is more likely to be on improved health-care processes than improved patient outcomes” (Hanney et al. 2013, p. 83). In line with this statement, the data of this thesis showed better adherence to treatment guidelines in hospitals participating in clinical trials as opposed to non-trial hospitals across all QS-OVAR cohorts. However, the observed effects did not consistently translate into survival differences and the beneficial effect on survival observed in the first cohort was not replicated in the subsequent two cohorts.

Finally, it should be mentioned again that even though, the adherence to treatment improved over time, the overall survival outcomes did not change much from the first to the last QS-OVAR cohort. This is in line with the general observation that despite increased awareness of the disease, curative and survival trends in ovarian cancer have not substantially improved over the last decades. This is primarily due to the ongoing challenge of early diagnosis, with treatment playing a secondary role. In addition, and unfortunately, for the patients, many factors may influence survival. Several factors have been explored as predictors of not receiving recommended comprehensive surgery in ovarian cancer (e.g. Goff et al. 2007). In addition, even the administration of a stage-appropriate chemotherapy does not guarantee that the patient will respond to the administered treatment. Chemosensitivity of the tumor can hardly be influenced and promising chemosensitivity assays to determine the drug sensitivity of solid tumors have been used only since a decade (Lee et al. 2014). Besides that, survival can be impacted by subsequent therapies (e.g., interval debulking, second line chemotherapy) that were not considered in this analysis. Finally, a general observation is that despite the enormous interest in treatment breakthroughs in oncology, the advances taking place are rather modest and only a very small number of new treatments make it from research to medical practice (Saltz 2008).

4.4 Challenges

The present thesis investigated whether research-active health care institutions deliver better care and outcomes than their less engaged peers do. This question has superseded the traditionally investigated question of whether outcomes of patients participating in trials differ from patients who do not participate in trials. Whereas the shortcomings of the latter question have been discussed in the introduction already, the researcher interested in answering the question about effects of clinical research on institutional level is confronted with numerous methodological challenges and practical issues as well (Pater et al. 2011). Addressing these methodological challenges and practical issues is crucial for accurately assessing the impact of research activity on healthcare institutions, and ultimately, for understanding how research engagement can lead to better patient care and outcomes.

First, the choice of the adequate study design including an appropriate control group is not as easy as it seems at first glance. Braunholtz et al. (2001) already suggested a less biased analysis by comparing trial patients with patients of non-recruiting physicians or institutions. In a similar vein, Peppercorn et al. (2004) proposed prospective cohorts of patients at institutions not participating in the trial as controls. A major problem is that providers not involved in research usually do not provide data that would serve as useful control group. But without this data, it becomes challenging to validate the effectiveness of clinical research, eliminate potential biases, and ensure the generalizability of the findings. But even in the research-active hospitals, good-quality data on the whole system rather than just the description of those patients that are actually involved in trials is required meaning that data from all patients of the research-active as well as the research-inactive institutions have to be collected. In an editorial about center size and the quality of cancer treatment, Corry et al. (2015) stated that it would be useful to obtain the total number of cancer patients seen in centers with expertise, as opposed to only those patients enrolled in trials. According to the authors, this would enable solid recommendations for the minimum number of patients per center that is required to achieve optimal survival outcome. Studies investigating the impact of institutional or structural characteristics must, therefore, ensure that all patients with the disease being studied are included and followed up for a sufficient duration of time within the compared institutions. It does not come as a surprise that such studies are rather uncommon because of their logistic difficulty and the limited opportunities for proper prospective studies. In practice, follow-up is rather short and data are often collected retrospectively.

I was fortunate to have access to data of such high quality. In the first phase of the OS-OVAR study, all German gynecology departments were contacted and requested to report the number of newly diagnosed patients with ovarian cancer. Approximately one year later, the second phase was initiated. At this point, all hospitals that responded during the first phase were asked to document all patients diagnosed in the third quarter of that year. This situation is unique, as typically data is only available from patients participating in trials. In this case, data was obtained not only from hospitals not participating in trials but also from non-trial patients within trial hospitals. Trained data managers of QS-OVAR cross-checked the report forms with surgical and pathologic reports. The hospital declarations regarding participation in clinical trials of one of the two study groups were double-checked as well. The available data included all relevant patient and disease characteristics, even comorbidities were collected. Survival information was collected for at least three years after diagnosis in the first cohort and at least four years in the later cohorts which is considered sufficient at least for patients with advanced ovarian cancer. In addition, the effect of study participation was deemed to cover more than only recruitment of some patients in the three months study period of the respective year. Study participation was also counted if no patient was enrolled in trials within the study period but if patients had been recruited before the third quarter. For example, QS-OVAR 2001 included more than one third of all patients diagnosed with ovarian cancer within the observation period even though these patients were reported from only around 15% of all German hospitals. This could, however, already reflect centralization. Nevertheless, the data from QS-OVAR was the most representative data available as national cancer registries were not established in Germany at that time.

The proper definition of “research activity” is also critical for investigating the question of this thesis. Even with a careful definition of research activity before the start of any evaluation, it is obvious that the concept is complex and includes more aspects than the simplistic operational dichotomy as used here (participation in AGO-OVAR or NOGGO trials in the respective year). The available literature on the relationship between research participation at the institutional level and health care outcomes is promising but limited, and the mechanisms by which research activity may lead to improved outcomes are not well understood. To examine the effects and to explore the corresponding mechanisms, the use of a conceptual framework such as Donabedian’s health care triad of structure, process and outcome has been therefore recommended in other indications (e.g., Lawson and Yazdany 2012). A unique feature of the present thesis is the integration of the three components within a single mediation analysis with survival as outcome and two mediators, taking into account important patient as well as disease characteristics. The latter reflects a significant extension of Donabedian’s simple framework and is in line with more recent recommendations (e.g., Selby at al. 2019a; b).

The complexity of the underlying mechanisms in research activity and health outcomes has led to a focus on not only whether an association exists ("Does it work?") but also on understanding the causal pathways for better outcomes in research-active hospitals. This is crucial for identifying processes to target for quality improvement. However, previous studies have primarily focused on the pragmatic question of research activity's effect on patient outcomes, with little attention paid to explaining these effects ("How does it work?"). In accordance with this, Boaz et al. (2015, p. 13), concluded: “However, although the focused review also identified a range of mechanisms through which engagement by clinicians and health care organizations in research might result in improved health care performance, and the wider review added additional evidence, it remains unclear how these effects are produced.” The causal mediation methodology implemented in the context of this thesis offers a way to approach this question.

Final challenge of the present thesis was the outcome variable, namely patient survival. With regards to efficacy, regulators request substantial evidence of clinical benefit from well-controlled trials demonstrating either prolonging patient survival or improving quality of life, or both. Especially in oncology, increasing survival and improving quality of life are the most important therapeutic objectives for many patients (Ocana and Tannock 2011). Overall survival is unambiguous, less subject to interpretation bias and represents a concrete direct benefit to the patient. In line with this, overall survival has been the most accepted outcome in ovarian cancer research (Herzog et al. 2014). In the most recent GCIG recommendations, overall survival remains an ideal primary endpoint for first-line trials. Nevertheless, it has been acknowledged that it is difficult to demonstrate a survival benefit in ovarian cancer because of long post progression survival and potential cross-over (Karam et al. 2017). In addition, in order to observe benefits in objective or so-called "hard" patient outcomes such as overall survival, the underlying effects might need to be much stronger. It has even been shown that survival is not always closely linked to the quality of care (Landon et al. 1998; Landon et al. 2001), in part because it is affected by other factors that are not related to the treatment provided. And even if this is the case, the question remains how much of the effect is mediated by actual treatment and how much is caused by other potentially relevant factors. Consequently, studies that evaluate how research activity affects quality of care need to gather all possible information on institutional infrastructure, processes and outcomes of care, patient and disease characteristics. The latter are needed to adjust for differences in case mix when exploring if and why outcomes may differ in relation to all the possible structural factors (Krzyzanowska et al. 2011). This was done in this thesis to the maximum possible extent.

4.5 Limitations

This study has several limitations. First, it might be argued that the sole participation in a quality assurance program and response to the survey are indicators of research activity. As the visibility and the reputation of the QS-OVAR program increased over the years, this might have led to a general higher hospital participation in the later cohorts. The increased popularity might also have led to a systematic dropout of low-performing hospitals. Besides that, some confounding between institutional research participation and patient volume as well as level of care was apparent at hospital level. Hospital volume as measured by the number of patients treated per year and the hospital level of care might both have an impact on patient outcomes. On the other hand, the data analyzed in this thesis did not support this assumption. Second, hospitals were categorized as trial hospitals and non-trial hospitals based on their participation in clinical trials of two German study groups. However, the non-trial institutions might have participated in other type of research. Unfortunately, this data was not collected. Inclusion of these institutions may have diluted the effect of institutional research participation on clinical outcomes, especially in the later cohorts.

The analyses were adjusted for case-mix, including both demographic and disease characteristics, which is essential in any study evaluating differences in outcomes between institutions. However, this is an observational study and the possibility of residual confounding must be acknowledged. The only way to avoid this type of confounding would be randomization. However, randomized trials in this area are hard to imagine. Randomizing patients to research-active versus inactive institutions might be considered unethical and recruitment of patients as well as hospitals would be rather challenging; not to mention required funding and policy implications. Lange et al.'s (2012; 2014) approach requires that there is no unmeasured confounding for the exposure-outcome relationship, exposure-mediator relationship, and mediator-outcome relationship. A limitation to the present analysis is that, for example, no data on socioeconomic status was available from the quality assurance program. Patients with high

socioeconomic status (e.g., higher income and insurance status) may find it easier and hence choose to travel to specialized hospitals, which may be more likely to be research-active. In contrast, patients with a low socioeconomic status may be less likely to choose research-active hospitals. The total effect of research activity would then at least be partially caused by better general survival in patients with a higher socioeconomic status (i.e., unmeasured confounding in the direct effect between exposure and outcome). Lower socioeconomic status has occasionally been reported to be associated with lower likelihood of receiving appropriate therapy (i.e., unmeasured confounding in the exposure-mediator relationship). Although these limitations formally apply to the present analysis, the current evidence for treatment differences and prognostic value of socioeconomic status is rather inconsistent. While some authors report that adherence to treatment guidelines for advanced-stage ovarian cancer is associated with equivalent survival benefit across socioeconomic subgroups (Bristow et al. 2015b), others reported that higher socioeconomic status was associated with a greater probability of undergoing surgical resection and with improved survival in patients with ovarian cancer (Gardy et al. 2019). In addition, this thesis used data from a German quality program. Germany has one of the best and most comprehensive health care systems in the world with an excellent insurance coverage.

Finally, the focus on only two mediators might be seen as a limitation regarding other possible causal pathways; it does, however, not invalidate or bias the effect estimates for the potential mediators (Lange et al. 2014). While it is crucial to include all relevant confounders, the causal models of Lange et al. (2012; 2014) do not require inclusion of all relevant mediators. Of course, the omission of relevant mediators might result in a higher weight of the direct path that summarizes all the unexplained effects between exposure and outcome. In fact, it is a plausible notion that there is no direct relationship between institutional research activity and patient survival at all; in other words, any effect of institutional research activity on patient survival is "mediated" by something in some way or another. The question here is again, what is the exact definition of such a structural characteristic like "research activity" and if this characteristic is a stand-alone feature or if this is rather a result of other factors.

4.6 Recent developments

4.6.1 Institutional research activity

In January 2023, the Lancet Oncology European Groundshot Commission published 12 recommendations, which, if acted upon, would reimagine a cancer research agenda for Europe (Lawler et al. 2023). Already in the abstract, the authors state that "patients treated in research-active hospitals have better outcomes than patients who are not treated in these settings" (p. e11). The commission unites a large and diverse group of specialists, accompanied by comprehensive new data regarding cancer research endeavors throughout Europe over the past 12 years. Their conclusion states that cancer research is vital for improved care, and research participation can boost organizational performance. They suggest that research-active hospitals may adopt new evidence-based practices more quickly. Furthermore, they contend that positive outcomes are directly linked to research activity, but that such a research activity needs to be broad and needs to cover domains from public health and cancer through to surgery, chemotherapy and radiotherapy, and even palliative care. As a result, a research-active health system that supports a wide variety of fundamental research, discovery research, and applied cancer research (including the translation of this research into patient-focused applications) is essential for truly improving patient outcomes. The commission's recommendations build upon some of the research findings cited in this thesis, especially the more recent studies exploring the relationship between research activity and improved patient outcomes (with more detailed results available in Section 4.3): For example, Boaz et al. (2015) concluded that clinicians and healthcare organizations

involved in research are likely to see improvements in healthcare performance, even if that was not the primary research goal. They also noted that traditional performance measures often overlooked research activity. Similarly, Harding et al. (2017) found a positive association between "research culture" and organizational performance, concluding that a stronger research culture benefits patients, staff, and the organization. Downing et al. (2017) reported that patients with colorectal cancer treated in research-active hospitals had better outcomes, with the impact of research participation on survival being comparable to the overall patient population impact seen after a successful interventional trial. Selby et al. (2019a; b) explored how clinical cancer research can enhance health outcomes, concluding that institutions actively engaged in research are generally more receptive to innovation and incorporate them at a faster pace. The authors emphasized the importance of evidence-based practice derived from the current literature, which serves as the foundation for guidelines and best practice in trials. They also highlighted the crucial role of discovery science and translational research in driving innovations tested in RCTs. When these trials yield positive results, they create a new evidence base for best practices, ultimately leading to improved patient outcomes both within the trial and for the larger patient population benefiting from the adoption of novel approaches. The authors suggested that these mechanisms collectively contribute to better patient outcomes through clinical research and advocated for investing in the necessary infrastructure as a potentially cost-effective approach to improving healthcare outcomes in oncology. Overall, clinical trial participation's impact on cancer patient outcomes remains relevant yet a subject of debate (Engelbak Nielsen et al. 2020).

4.6.2 Ovarian cancer treatment and outcomes in Germany

The data from the ongoing quality assurance program QS-OVAR has consistently shown that the quality of ovarian cancer treatment in Germany has steadily improved over the years (Harter et al. 2020). The latest 2016 cohort results with a 5-year follow-up were presented at several national and international conferences in 2022. Unsurprisingly, they reaffirmed that the outcome of patients with ovarian cancer still largely depends on the quality of treatment and the expertise of the treating physicians and institutions. In early ovarian cancer, the best prognosis could be achieved if surgery and chemotherapy were done according to treatment guidelines (Sehouli et al. 2022). In advanced ovarian cancer, the results confirmed that overall survival was best when complete tumor resection was achieved at primary surgery and patients received combination chemotherapy with maintenance treatment (Mahner et al. 2022). These findings further support the importance of research activity and adherence to evidence-based practices in improving patient outcomes. The enduring relevance of the topic today is a testament to the insights gained from the past cohorts of QS-OVAR.

4.6.3 Mediation analysis

The concept of mediation analysis is not new, as it can be traced back to Wright (1934) and gained popularity in the social sciences almost four decades ago with the publication of the influential work by Baron and Kenny (1986). Nonetheless, the mediation analysis techniques remain an active research area, with growing interest in causal processes and ongoing advancements in the theory, methods, and computing tools (VanderWeele 2015; VanderWeele 2016). Mediation analysis, once predominantly utilized in psychology, is now increasingly embraced in healthcare research where it serves two main objectives: understanding the mechanisms underlying the treatment effects and identifying potential targets for upcoming interventions. Most importantly, it has been recognized that mediation analyses of randomized trials and observational studies can generate evidence about the mechanisms by which interventions and exposures may influence health outcomes (Emsley et al. 2010). The hope is that a deeper understanding of these mechanisms will allow for their specific improvement,

eventually leading to improved health outcomes. Therefore, it does not come as a surprise that the number of publications that include the term "mediation analysis" in the title or text has been steadily increasing over the past few years (Nguyen et al. 2021). To date (Google Scholar, August 2023), Baron and Kenny's (1986) seminal paper on mediation analysis has been cited over 120,000 times. This publication seems to have sparked most of the mediation analyses that are still seen today.

Although the "traditional" mediation approach as proposed by Baron and Kenny in 1986 is intuitive and easy to implement, it has several limitations (see also Section 1.4 of this thesis). First, it requires fulfilling a series of stepwise tests to quantify the degree of mediation starting with the exposure-outcome effect. The proposed tests are conservative and have low statistical power due to the condition of requiring a non-zero relationship between exposure and outcome for further investigation of mediation (MacKinnon et al. 2002). The evidence for mediation is strongest when there is full mediation meaning that there is an indirect effect of exposure on outcome via the mediator but no direct effect. There is only partial mediation when both indirect and direct effects are present. Iacobucci (2008, p. 12) noted that, "when all tests are properly conducted and reported, the majority of articles conclude with partial mediation." Finally, the effect definitions are highly model dependent. To quantify the degree of mediation, simple formulas combine parameter estimates obtained from a series of regressions. The resulting difference and product tests were originally intended for linear relationships with continuous outcomes. Unfortunately, these traditional methods do not work for non-normally distributed variables and especially not for censored time-to-event outcomes (Gelfand et al. 2016).

Fortunately, over the past decade, mediation analysis has witnessed rapid and significant methodological advancements overcoming the constraints of the traditional approach. In this context, especially the incorporation of the causal inference in mediation analysis can be considered a game changer as described by Nguyen et al. (2021): First, causal mediation analysis enabled effect definitions with more specific causal interpretation than available before. Moreover, causal mediation analysis clarified the assumptions required for the effect identification, and eventually broadened the spectrum of options for estimating such effects from data. Critically, the causal inference method distinguishes between the definition of an effect that researchers aim to estimate and the process of estimating it. Effects are defined in a model-independent way, truly relying on reasoning that aligns with the concept of a causal effect. The framework for this objective is the counterfactual framework (Rubin 2004), where a causal effect is described as a comparison between potential outcomes under two distinct conditions for the same individual or group.

At least five reviews on the use of mediation analyses were published in the most recent years. The first review focused on mediation models with time-to-event outcomes (Lapointe-Shaw et al. 2018). Two more reviews focused either on mediation analysis of experimental data from randomized controlled trials (Vo et al. 2020) or mediation analysis of observational data (Rijnhart et al. 2021). Another review focused on a special area of interest, namely the use of mediation analysis in the field of psychology and psychiatry (Stuart et al. 2021). Finally, one "meta review" assessed other systematic reviews of mediation analysis (Cashin et al. 2019). All reviews distinguished between the two approaches for conducting mediation analysis: The traditional approach based on the comparison between regression models with and without conditioning on the mediator and the more modern approach using the counterfactual framework decomposing the total effect into direct and indirect effects. In the following paragraph, I will provide a concise summary of the five recent reviews on mediation analyses, which cover different focuses but share similar conclusions on mediation analysis usage and reporting.

Cashin et al. (2019) found in their review of reviews that over the past decade, there were 54 published systematic reviews including 2008 primary mediation studies, across 11 health care fields and 26 health conditions. Limitations in reporting quality were identified in 18 (33%) of these 54 systematic reviews , and 29 (54%) systematic reviews stated limitations in the methodological conduct of the primary mediation studies. The authors concluded that there is an urgent need to improve the reporting and methodological quality of studies investigating mediation. This conclusion underscores the importance of rigorous methodology and transparent reporting in mediation studies, a sentiment echoed by other researchers in the field.

Lapointe-Shaw et al. (2018) studied the use and reporting of mediation analysis with a time-to-event outcome in healthcare research. This review consisted of 149 primary studies, published from 1997 to 2016, including the main publication of the present thesis (Rochon et al. 2014). The review found an increasing use of mediation analysis across various clinical fields over time, with a focus on understanding patient-related factors that explain the onset of a disease. Most (74%) of the 149 studies were published after 2011, and the annual number of studies nearly doubled in the last year (from 21 to 40 studies per year). Most studies used traditional mediation analysis, and even more than half of the studies (58%) took a traditional approach using the four steps proposed by Baron and Kenny (1986) for the evaluation of mediation by means of regression models (see also Section 1.4 of this thesis). Only 35% of the reviewed studies mentioned any of the underlying assumptions or limitations fundamental to a causal interpretation of mediation analysis. Most studies (77%) used a Cox proportional hazards regression for the outcome of interest. The authors highlighted the need for better reporting, consideration of assumptions, and awareness of the limitations of Cox proportional hazards regression for common (i.e., not rare) outcomes in mediation analysis (VanderWeele, 2011).

Vo et al. (2020) conducted a systematic review by searching MEDLINE (from January 2017 until December 2018) for RCTs or secondary analyses of previously published RCTs that reported a mediation analysis. In line with Lapointe-Shaw et al. (2018), the authors found that most mediation analyses used the traditional Barron and Kenny (1986) stepwise approach (96%). In contrast, the causal inference approach, was rarely used in clinical research practice. Most studies did not report any sample size calculations for the mediation analysis, nor did they assess potential treatment-by-mediator interactions. In 53% of the investigated studies, mediators and outcomes were measured simultaneously, raising into question if the authors adequately included causal relationships in their theoretical considerations. In 57% of studies, mediator-mediator and mediator-outcome confounders were adjusted for in the analysis, although adjustment was often limited to few potential confounders. Only about 30% of studies discussed the assumptions underlying the conducted mediation analysis.

Rijnhart et al. (2021) included 174 studies in their review of mediation analysis in observational studies, most of which (71%) again applied traditional mediation analysis methods. Causal mediation analysis was mainly used to analyze more complicated mediation models, such as multiple mediator models. Most studies adjusted their analyses for measured confounders but did not perform any sensitivity analyses for unmeasured confounders nor did they assess the presence of an exposure-mediator interaction. The authors recommended that researchers employ causal mediation analysis and evaluate the plausibility of causal assumptions to guarantee the causal interpretation of direct and indirect effect estimates. Additionally, they suggested that the adoption of causal mediation analysis could be improved through instructional papers and the creation of software packages that simplify the estimation of causal effects for more complex models, such as multilevel and longitudinal mediation models.

Finally, Stuart et al. (2021) found in their review of 206 articles that included a mediation analysis published in top-tier psychology and psychiatry journals in 2013–2018, that nearly all the articles (97%) used the traditional mediation analysis approach and only 3% used causal mediation analysis. In most studies, the underlying assumptions were not tested or even mentioned. One important takeaway from this review was that mediation analysis inherently assumes that the exposure affects the mediator, and both the exposure and mediator influence the outcome. In other words, mediation is based on a causal hypothesis, which serves as the foundation for all methodological approaches used to evaluate it. This causal hypothesis necessitates a temporal sequence of these variables, which should be evident in the data. Without this temporal sequence in the data, it is impossible to confirm that the order is exposure, followed by mediator, and then outcome, rather than a different sequence. The authors discovered that only 24% of the 206 articles reviewed had a temporal ordering of all three variables: exposure, mediator, and outcome. This percentage is similar to what was found in a 2007 review of psychology research (Maxwell and Cole 2007). Temporal ordering was more prevalent in randomized trials, with approximately 80% having at least temporal ordering of exposure before mediator or mediator before outcome (and nearly 50% having full temporal ordering of all three variables); this finding aligns with the results reported by Vo et al. (2020). Nonetheless, one should keep in mind that a given temporal order of *A*, *M*, and *Y* does not automatically imply a causal relationship. Moreover, obtaining cross-sectional data does not rule out causal hypotheses and testing for mediation may still be suitable using purely retrospective data collected at a single time point.

In conclusion, the five reviews consistently highlighted that while the use of mediation analysis has increased, the uptake of causal mediation analysis in applied research remains relatively low and the quality of reporting in these studies is poor. To overcome the latter problem, the US Berkeley Initiative for Transparency in the Social Sciences and the Center for Effective Global Action funded a project to develop "A Guideline for Reporting Mediation Analyses": AGReMA. The international consensus process described in Cashin et al. (2020) produced a 25-item AGReMA checklist statement and a 9-item AGReMA short-form (AGReMA-SF). The scope of the AGReMA statement covers primary and secondary mediation analyses of randomized trials and observational studies, and it is intended to be general so that it can guide the reporting of most mediation analyses (Lee et al. 2021). Through implementation of this guideline, causal mediation analysis may gain momentum, and researchers may become more inclined to adopt these methods in their future studies. Overall, the expectation is that AGReMA will help to improve the transparency, reproducibility, and overall quality of mediation analysis reporting, contributing to more robust and reliable findings, and ultimately enhancing the impact of mediation analysis in healthcare research.

Key technical articles on causal mediation analysis were published between 2009 and 2012 (VanderWeele and Vansteelandt 2009; Imai et al. 2010; VanderWeele and Vansteelandt 2010; Pearl 2012; Lange et al. 2012; Vansteelandt et al. 2012), but its application has been slow due to its novelty and technical complexity. The lack of practical guidelines and understanding of causal approaches, as well as unfamiliarity with the necessary coding, have also been barriers. Beyond the anticipated positive effects of AGReMA, the growing availability of causal mediation analysis software is expected to increase its use. Recent guides such as VanderWeele's (2016) practitioner's guide or Lipsky and Greenland's (2022) guide illustrating the use of causal directed graphs are helping to facilitate learning and application of this methodology in a broader community. Other articles like Preacher's (2015) review synthesizing four methodological research areas or the most recent work by Cashin et al. (2023) focusing on the "understanding how health interventions or exposures produce their effects using

mediation analysis" (p. 1), are aiming to disseminate methodological advancements not only to statisticians and suggest future research directions for mediation analysis. Web applications for power and sample size calculations for causal mediation analysis have just been made available (e.g., Qin, in press). In addition, prominent research centers and funding organizations have recently advocated for mediation analysis. For example, the new version of the Medical Research Council framework for developing and evaluating complex interventions not only encourages "a wider range and combination of research perspectives and methods, which answer questions beyond efficacy and effectiveness" but even explicitly uses the term mediation in one of the questions that should be asked by decision makers: "How are the intervention effects mediated by different settings and contexts?" (Skivington et al. 2021, p. 3). Finally, there is also growing interest from regulators and the pharmaceutical industry in statistical analyses that provide insights into causal mechanisms of action. A recent example is the causal mediation analysis with time-to-event endpoints (Vansteelandt et al. 2019) based on data from LEADER, a large cardiovascular outcome trial (Marso et al. 2016).

The rising interest in causal mediation analysis from various stakeholders, including regulators, payers, research institutions and pharmaceutical sponsors, indicates a promising future for this method. As understanding and resources grow, its use in healthcare research will likely increase, improving the reliability of findings and offering insights into causal mechanisms. Harnessing the full potential of this innovative approach will hopefully drive progress in medicine and lead to better patient outcomes.

4.7 Conclusion

In conclusion, this thesis explored the widely held assumption that engagement in research by clinicians and health care institutions improves health care performance and outcomes. The focus was specifically on the treatment of ovarian cancer in Germany. The study uncovered considerable deficiencies in treatment and identified areas for improvement. A beneficial trial participation effect was found in the first cohort, suggesting that hospital participation in trials may influence treatment processes and patient survival. However, this hypothesis was not confirmed in later cohorts, leaving the impact of clinical trial participation on cancer patient outcomes a controversial topic.

Improving the quality of care is crucial, and based on the results from this thesis, the priority should be first on optimizing processes of care (i.e., treatment) and then on enhancing outcomes (i.e., patient survival). The effects of institutional research activity on patient outcomes remain not fully understood, emphasizing the need for further work to investigate causal pathways. This thesis recommends employing causal mediation analysis to gain insights into the mechanisms of exposure-outcome effects. The introduced methodology enables researchers to explore potential mechanisms, shifting from the question "Does it work?" to answering "How does it work?" for a wide range of variable types.

5 Summary

5.1 Summary (English)

Randomized controlled trials are considered the gold standard for the development and approval of new treatments in evidence-based medicine. The effects of individual participation in clinical research have been discussed in the literature for decades, particularly in the field of oncology. The present thesis investigated whether institutional participation in clinical trials has an influence on the quality of care and how this affects treatment outcomes in patients with ovarian cancer, regardless of whether the patients participate in clinical trials or not. The focus of the thesis was on exploring the underlying mechanisms in Donabedian's triad of health care with structure, process, and outcome as a conceptual framework, while taking into account methodological challenges of observational studies.

The thesis used data from three cohorts of a German quality assurance program for ovarian cancer. The hypothesis was that institutional research activity is associated with better adherence to treatment guidelines for ovarian cancer and that this, in turn, leads to better treatment outcomes for the affected patients. Surgical standards and guideline-based administration of chemotherapy were assumed as potential mediators between hospital-related participation in clinical trials and overall survival of patients. In the study, a new method for causal mediation analysis was introduced and further developed for various scenarios and types of variables, enabling the investigation of direct and indirect effects of institutional research activity on patient survival, i.e., a time-to-event outcome with censoring. This made it possible for the very first time in this context to examine and quantify causal pathways between structural characteristics of the treating institution, process variables reflecting treatment quality, and patient-relevant outcomes such as survival. This also allowed for the practical embedding of Donabedian's quality model and the integration of its three components (structure, process, and outcome) into a single analysis model. In addition, patient and disease characteristics could be taken into account. The problem of clustering was solved by bootstrap techniques.

The study showed heterogeneity in the treatment of ovarian cancer in Germany and identified opportunities for improvement. It turned out that hospitals participating in clinical trials provided better care, at least in the first cohort, and thus achieved longer overall survival for their patients. In later cohorts, however, this effect was no longer observed. Therefore, the dissertation concluded that the question of whether participation in clinical trials improves the outcome in patients with cancer diagnosis remains highly relevant but controversial.

Adherence to treatment guidelines for ovarian cancer improved over time. Surprisingly, this improvement in treatment quality did not translate into even longer survival for the affected patients. Future efforts in this area should therefore be aimed at translating new findings from clinical research into daily practice in a timely manner, but also at further examining influencing variables on treatment success in order to ensure the best possible care for patients with ovarian cancer while at the same time achieving longer overall survival. For the investigation of the underlying mechanisms of action, causal mediation analyses should be used.

5.2 Zusammenfassung (Deutsch)

Randomisierte kontrollierte Studien gelten als Goldstandard für die Entwicklung und Zulassung neuer Behandlungsmethoden in der evidenzbasierten Medizin. Effekte individueller Teilnahme an klinischer Forschung werden in der Literatur seit Jahrzehnten diskutiert, insbesondere im Bereich der Onkologie. In der vorliegenden Arbeit wurde untersucht, ob institutionelle Teilnahme an klinischen Studien einen Einfluss auf die Versorgungsqualität hat, und wie sich diese auf Behandlungsergebnisse bei Patientinnen mit Eierstockkrebs auswirkt – unabhängig davon, ob die Patientinnen selbst an klinischen Studien teilnehmen oder nicht. Im Mittelpunkt der Arbeit lag die Erforschung der Wirkmechanismen in Donabedians sogenannter Triade der Gesundheitsversorgung mit Struktur, Prozess und Ergebnis als konzeptionellem Rahmen, und unter Berücksichtigung methodischer Herausforderungen von Beobachtungsstudien.

Für die Arbeit wurden Daten aus drei Kohorten eines deutschen Qualitätssicherungsprogramms der Kommission Ovar der Arbeitsgemeinschaft Gynäkologische Onkologie verwendet. Die Hypothese war, dass institutionelle Forschungsaktivität mit besserem Befolgen von Behandlungsleitlinien einhergeht und dies wiederum zu besseren Behandlungsergebnissen bei Patientinnen mit Eierstockkrebs führt. Als potenziell vermittelnde Variablen wurden chirurgische Standards und leitliniengerechte Chemotherapie angenommen. Es wurde eine neue Methode zur kausalen Mediationsanalyse eingeführt und für verschiedene Szenarien und Arten von Variablen weiterentwickelt, mit der direkte und indirekte Auswirkungen institutioneller Forschungsaktivität auf das Überleben von Patientinnen, also einer zeitlichen Zielgröße mit Zensierung, untersucht werden können. Hierdurch war es in diesem Kontext erstmals möglich, kausale Pfade zwischen Struktureigenschaften der behandelnden medizinischen Einrichtung, über Prozessvariablen, die die Behandlungsqualität widerspiegeln, hin zu patientenrelevanten Zielgrößen wie Überleben zu prüfen und zu quantifizieren. Dies erlaubte zudem die praktische Einbettung des Qualitätsmodells nach Donabedian und die Integration seiner drei Komponenten (Struktur, Prozess und Ergebnis) in ein einziges Analysemodell. Außerdem konnten Patientinnen- und Krankheitsmerkmale berücksichtigt werden. Das Problem der Clusterung von Patientinnen in behandelnden Krankenhäusern wurde durch Bootstrap-Techniken gelöst.

Die Studie zeigte die Heterogenität in der Behandlung von Eierstockkrebs in Deutschland und identifizierte Möglichkeiten für Verbesserungen. Es stellte sich heraus, dass Krankenhäuser, die an klinischen Studien teilnahmen, zumindest in der ersten Kohorte bessere Versorgung boten und hierdurch auch ein längeres Gesamtüberleben für ihre Patientinnen erzielten. In späteren Kohorten war dieser Effekt allerdings nicht mehr zu beobachten. Daher kam die Dissertation zu dem Schluss, dass die Frage, ob die Teilnahme an klinischen Studien die Prognose bei einer Krebsdiagnose verbessert, zwar nach wie vor hohe Relevanz hat, allerdings kontrovers bleibt.

Die Einhaltung von Behandlungsleitlinien bei Eierstockkrebs verbesserte sich im Laufe der Zeit. Überraschenderweise übersetzte sich diese Verbesserung der Behandlungsqualität jedoch nicht in noch längere Überlebenszeiten für die betroffenen Patientinnen. Zukünftige Bemühungen in diesem Bereich sollten daher darauf ausgerichtet sein, neue Erkenntnisse aus klinischer Forschung zeitnah in die tägliche Praxis zu übersetzen, aber auch weitere Einflussvariablen auf die Prognose näher zu untersuchen, um einerseits für Krebspatientinnen die bestmögliche Versorgung zu gewährleisten, zugleich aber auch ein längeres Gesamtüberleben zu erzielen. Für die Untersuchung der zugrundeliegenden Wirkmechanismen sollten kausale Mediationsanalysen verwendet werden.

6 Bibliography

American Cancer Society (2023). *Cancer Facts & Figures 2023.* Atlanta, Ga: American Cancer Society.

Andersen, M., Kragstrup, J., and Søndergaard, J. (2006). How conducting a clinical trial affects physicians' guideline adherence and drug preferences. *Journal of the American Medical Association*, *295*(23), 2759–2764.

Armstrong, D. K., Alvarez, R. D., Bakkum-Gamez, J. N., Barroilhet, L., Behbakht, K., Berchuck, A., Berek, J. S., Chen, L. M., Cristea, M., DeRosa, M., ElNaggar, A. C., Gershenson, D. M., Gray, H. J., Hakam, A., Jain, A., Johnston, C., Leath, C.A. III, Liu, J., Mahdi, H., Matei, D., McHale, M., McLean, K., O'Malley, D. M., Penson, R. T., Percac-Lima, S., Ratner, E., Remmenga, S. W., Sabbatini, P., Werner, T. L., Zsiros, E., Burns, J. L., and Engh, A. M. (2019). NCCN Guidelines Insights: Ovarian Cancer, Version 1.2019. *Journal of the National Comprehensive Cancer Network, 17*(8), 896–909.

Ashcroft, R. (1999). Equipoise, knowledge and ethics in clinical research and practice. *Bioethics*, *13*, 314–326.

Barlin, J. N., Yu, C., Hill, E. K., Zivanovic, O., Kolev, V., Levine, D. A., Sonoda, Y., Abu-Rustum, N. R., Huh, J., Barakat, R. R., Kattan, M. W., and Chi, D. S. (2012). Nomogram for predicting 5-year disease-specific mortality after primary surgery for epithelial ovarian cancer. *Gynecologic Oncology, 125*(1), 25–30.

Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Bauknecht, T., Breitbach, G. P., and du Bois A (2000). Maligne Ovarialtumoren. In: Deutsche Krebsgesellschaft e.V. (Hrsg.), *Qualitätssicherung in der Onkologie. Kurzgefasste interdisziplinäre Leitlinien.* München: Zuckschwerdt, pp. 301–318.

Bertelsen, K. (1991). Protocol allocation and exclusion in two Danish randomised trials in ovarian cancer. *British Journal of Cancer*, *64*(6), 1172–1176.

Birkmeyer, J. D., Dimick, J. B., and Birkmeyer, N. J. (2004). Measuring the quality of surgical care: structure, process, or outcomes? *Journal of the American College of Surgeons*, *198*(4), 626–632.

Boaz, A., Hanney, S., Jones, T., and Soper, B. (2015). Does the engagement of clinicians and organisations in research improve health care performance: a three-stage review. *BMJ Open, 5,* e009415.

Boyd, L. R., and Muggia, F. M. (2018). Carboplatin/Paclitaxel induction in ovarian cancer: the finer points. *Oncology*, *32*(8), 418–424.

Braunholtz, D. A., Edwards, S. J., and Lilford, R. J. (2001). Are randomized clinical trials good for us (in the short term)? Evidence for a "trial effect". *Journal of Clinical Epidemiology*, *54*(3), 217–224.

Bray, F., Laversanne, M., Weiderpass, E., and Soerjomataram, I. (2021). The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer, 127*(16), 3029–3030.

Breyer, J. Z., Giacomazzi, J., Kuhmmer, R., Lima, K. M., Hammes, L. S., Ribeiro, R. A., Kops, N. L., Falavigna, M., and Wendland, E. M. (2019). Hospital quality indicators: a systematic review. *International Journal of Health Care Quality Assurance, 32*(2), 474–487.

Bristow, R. E., Tomacruz, R. S., Armstrong, D. K., Trimble, E. L., and Montz, F. J. (2002). Survival effect of maximal cytoreductive surgery for advanced ovarian carcinoma during the platinum era: a meta-analysis. *Journal of Clinical Oncology*, *20*(5), 1248–1259.

Bristow, R. E., Chang, J., Ziogas, A., and Anton-Culver, H. (2013). Adherence to treatment guidelines for ovarian cancer as a measure of quality care. *Obstetrics and Gynecology*, *121*(6), 1226–1234.

Bristow, R. E., Chang J., Ziogas A., Campos B., Chavez L. R., and Anton-Culver, H. (2015a). Impact of National Cancer Institute Comprehensive Cancer Centers on ovarian cancer treatment and survival. *Journal of the American College of Surgeons, 220*(5), 940–950.

Bristow, R. E., Chang, J., Ziogas, A., Campos, B., Chavez, L. R., and Anton-Culver, H. (2015b). Sociodemographic disparities in advanced ovarian cancer survival and adherence to treatment guidelines. *Obstetrics and Gynecology, 125*(4), 833–842.

Cabasag, C. J., Arnold, M., Butler, J., Inoue, M., Trabert, B., Webb, P. M., Bray, F., and Soerjomataram, I. (2020a). The influence of birth cohort and calendar period on global trends in ovarian cancer incidence. *International Journal of Cancer, 146*(3), 749–758.

Cabasag, C. J., Butler, J., Arnold, M., Rutherford, M., Bardot, A., Ferlay, J., Morgan, E., Møller, B., Gavin, A., Norell, C. H., Harrison, S., Saint-Jacques, N., Eden, M., Rous, B., Nordin, A., Hanna, L., Kwon, J., Cohen, P. A., Altman, A. D., Shack, L., and Soerjomataram, I. (2020b). Exploring variations in ovarian cancer survival by age and stage (ICBP SurvMark-2): A population-based study. *Gynecologic Oncology, 157*(1), 234–244.

Cashin, A. G., Lee, H., Lamb, S. E., Hopewell, S., Mansell, G., Williams, C. M., Kamper, S., Henschke, N., and McAuley, J. H. (2019). An overview of systematic reviews found suboptimal reporting and methodological limitations of mediation studies investigating causal mechanisms. *Journal of Clinical Epidemiology, 111*, 60–68.

Cashin, A. G., McAuley, J. H., Lamb, S. E., Hopewell, S., Kamper, S. J., Williams, C. M., Henschke, N., and Lee, H. (2020). Development of a guideline for reporting mediation analyses (AGReMA). *BMC Medical Research Methodology, 20,* 19.

Cashin, A. G., McAuley, J. H., VanderWeele, T. J., and Lee, H. (2023). Understanding how health interventions or exposures produce their effects using mediation analysis. *The BMJ, 382*.

Chalmers, I., Bracken, M. B., Djulbegovic, B., Garattini, S., Grant, J., Gülmezoglu, A. M., Howells, D. W., Ioannidis, J. P., and Oliver, S. (2014). How to increase value and reduce waste when research priorities are set. *The Lancet*, *383*(9912), 156–165.

Chen, A. Y., Schrag, N., Hao, Y., Flanders, W. D., Kepner, J., Stewart, A., and Ward, E. (2006). Changes in treatment of advanced laryngeal cancer 1985–2001. *Otolaryngology–Head and Neck Surgery*, *135*(6), 831–837.

Chi, D. S., Palayekar, M. J., Sonoda, Y., Abu-Rustum, N. R., Awtrey, C. S., Huh, J., Eisenhauer, E. L., Barakat, R. R., and Kattan, M. W. (2008). Nomogram for survival after primary surgery for bulky stage IIIC ovarian carcinoma. *Gynecologic Oncology*, *108*(1), 191–194.

Clark, W. F., Garg, A. X., Blake, P. G., Rock, G. A., Heidenheim, A. P., and Sackett, D. L. (2003). Effect of awareness of a randomized controlled trial on use of experimental therapy. *Journal of the American Medical Association*, *290*(10), 1351–1355.

Clark, T. G., Stewart, M. E., Altman, D. G., Gabra, H., and Smyth, J. F. (2001). A prognostic model for ovarian cancer. *British Journal of Cancer*, *85*(7), 944–952.

Clark, T. G., Stewart, M., Rye, T., Smyth, J. F., and Gourley, C. (2007). Validation of a new prognostic index for advanced epithelial ovarian cancer: results from its application to a UK-based cohort. *Journal of Clinical Oncology*, *25*(35), 5669–5671. (Author reply 5670–5671).

Clarke, M., and Loudon, K. (2011). Effects on patients of their health care practitioner's or institution's participation in clinical trials: a systematic review. *Trials*, *12*, [16].

Coburn, S. B., Bray, F., Sherman, M. E., and Trabert, B. (2017). International patterns and trends in ovarian cancer incidence, overall and by histologic subtype. *International Journal of Cancer, 140*(11), 2451–2460.

Cole, S. R., and Hernán, M. A. (2002). Fallibility in estimating direct effects. *International Journal of Epidemiology*, *31*(1), 163–165.

Corry, J., Peters, L. J., and Rischin, D. (2015). Impact of center size and experience on outcomes in head and neck cancer. *Journal of Clinical Oncology*, *33*(2), 138–140.

Dalmartello, M., La Vecchia, C., Bertuccio, P., Boffetta, P., Levi, F., Negri, E., and Malvezzi, M. (2022). European cancer mortality predictions for the year 2022 with focus on ovarian cancer. *Annals of Oncology, 33*(3), 330–339.

De Angelis, R., Sant, M., Coleman, M. P., Francisci, S., Baili, P., Pierannunzio, D., Trama, A., Visser, O., Brenner, H., Ardanaz, E., Bielska-Lasota, M., Engholm, G., Nennecke, A., Siesling, S., Berrino, F., and Capocaccia, R. (2014). Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE – 5-a population-based study. *The Lancet Oncology*, *15*(1)*,* 23–34.

Dembo, A. J., Davy, M., Stenwig, A. E., Berle, E. J., Bush, R. S., and Kjorstad, K. (1990). Prognostic factors in patients with stage I epithelial ovarian cancer. *Obstetrics and Gynecology*, *75*(2), 263–273.

Donabedian, A. (1966). Evaluating the quality of medical care. *The Milbank Memorial Fund Quarterly*, *44*(3), 166–206.

Donabedian, A. (1980). *Explorations in quality assessment and monitoring: the definition of quality and approaches to its assessment.* Ann Arbor: Michigan Health Administration Press.

Donabedian, A. (1988). The quality of care. How can it be assessed? *Journal of the American Medical Association*, *260*(12), 1743–1748.

Downing, A., Morris, E. J., Corrigan, N., Sebag-Montefiore, D., Finan, P. J., Thomas, J. D., Chapman, M., Hamilton, R., Campbell, H., Cameron, D., Kaplan, R., Parmar, M., Stephens, R., Seymour, M., Gregory, W., and Selby, P. (2017). High hospital research participation and improved colorectal cancer survival outcomes: a population-based study. *Gut*, *66*(1), 89–96.

du Bois, A., Pfisterer, J., and Kellermann, L. (2001a). Die Therapie des fortgeschrittenen Ovarialkarzinoms in Deutschland: Ergebnisse einer Umfrage der Organkommission OVAR der Arbeitsgemeinschaft Gynäkologische Onkologie (AGO) in der Deutschen Krebsgesellschaft (DKG). *Der Gynäkologe, 11*(34), 1029–1040.

du Bois, A., Pfisterer, J., Kellermann, L., and Kreienberg, R. (2001b). Die Therapie des fortgeschrittenen Ovarialkarzinoms in Deutschland: Welchen Einfluss hat die Teilnahme an klinischen Studien? *Geburtshilfe und Frauenheilkunde*, *61*(11), 863–871.

du Bois, A., Quinn, M., Thigpen, T., Vermorken, J. B., Avall-Lundqvist, E., Bookman, M., Bowtell, D., Brady, M., Casado, A., Cervantes, A., Eisenhauer, E., Friedlaender, M., Fujiwara, K., Grenman, S., Guastalla, J. P., Harper, P., Hogberg, T., Kaye, S., Kitchener, H., Kristensen, G., Mannel, R., Meier, W., Miller, B., Neijt, J. P., Oza, A., Ozols, R., Parmar, M.K., Pecorelli, S., Pfisterer, J., Poveda, A., Provencher, D., Pujade-Lauraine, E., Randall, M., Rochon, J., Rustin, G., Sagae, S., Stehman, F., Stuart, G., Trimble, E., Vasey, P., Vergote, I., Verheijen, R., and Wagner, U. (2005b). 2004 consensus statements on the management of ovarian cancer: final document of the 3rd International Gynecologic Cancer Intergroup Ovarian Cancer Consensus Conference (GCIG OCCC 2004). *Annals of Oncology, 16*(Suppl. 8), viii7–viii12.

du Bois, A., Reuss, A., Pujade-Lauraine, E., Harter, P., Ray-Coquard, I., and Pfisterer, J. (2009a). Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: a combined exploratory analysis of 3 prospectively randomized phase 3 multicenter trials by the Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe Ovarialkarzinom (AGO-OVAR) and the Groupe d'Investigateurs Nationaux Pour les Études des Cancers de l'Ovaire (GINECO). *Cancer*, *115*(6), 1234–1244.

du Bois, A., Rochon, J., Lamparter, C., Pfisterer, J., and AGO Organkommission OVAR (2005a). Pattern of care and impact of participation in clinical studies on the outcome in ovarian cancer. *International Journal of Gynecological Cancer*, *15*(2), 183–191.

du Bois, A., Rochon, J., Pfisterer, J., and Hoskins, W. J. (2009b). Variations in institutional infrastructure, physician specialization and experience, and outcome in ovarian cancer: a systematic review. *Gynecologic Oncology*, *112*(2), 422–436.

Earle, C. C., Schrag, D., Neville, B. A., Yabroff, K. R., Topor, M., Fahey, A., Trimble, E. L., Bodurka, D. C., Bristow, R. E., Carney, M., and Warren, J. L. (2006). Effect of surgeon specialty on processes of care and outcomes for ovarian cancer patients. *Journal of the National Cancer Institute*, *98*(3), 172–180.

ECRI Institute (2002). Patients' reasons for participation in clinical trials and effect of trial participation on patient outcomes. *ECRI Health Technology Assessment Information Service, 74*, 1–38.

Emsley, R., Dunn, G., and White, I. R. (2010). Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. *Statistical Methods in Medical Research, 19*(3), 237–270.

Engelbak Nielsen, Z., Eriksson, S., Schram Harsløf, L. B., Petri, S., Helgesson, G., Mangset, M., and Godskesen, T. E. (2020). Are cancer patients better off if they participate in clinical trials? A mixed methods study. *BMC Cancer, 20,* [401].

European Society for Medical Oncology (ESMO) Guidelines Task Force (2001). ESMO minimum clinical recommendations for diagnosis, treatment and follow-up of ovarian cancer. *Annals of Oncology*, *12*(9), 1205–1207.

Ferlay, J., Ervik, M., Lam, F., Colombet, M., Mery, L., Piñeros, M., Znaor, A., Soerjomataram, I., and Bray, F. (2020a). Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. URL: https://gco.iarc.fr/today [accessed 30 August 2023].

Ferlay, J., Laversanne, M., Ervik, M., Lam, F., Colombet, M., Mery, L., Piñeros, M., Znaor, A., Soerjomataram, I., and Bray, F. (2020b). Global Cancer Observatory: Cancer Tomorrow. Lyon, France: International Agency for Research on Cancer. URL: https://gco.iarc.fr/tomorrow [accessed 30 August 2023].

Fielding, J. W., Fagg, S. L., Jones, B. G., Ellis, D., Hockey, M. S., Minawa, A., Brookes, V. S., Craven, J. L., Mason, M. C., Timothy, A., Waterhouse, J. A., and Wrigley, P. F. (1983). An interim report of a prospective, randomized, controlled study of adjuvant chemotherapy in operable gastric cancer: British Stomach Cancer Group. *World Journal of Surgery, 7*(3), 390–399.

Franke, R. H., and Kaul, J. D. (1978). The Hawthorne experiments: First statistical interpretation. *American Sociological Review*, *43*(5), 623–643.

Gardy, J., Dejardin O., Thobie A., Eid Y., Guizard A. V., and Launoy G. (2019). Impact of socioeconomic status on survival in patients with ovarian cancer. *International Journal of Gynecological Cancer, 29*(4), 792–801.

Gelfand, L. A., MacKinnon, D. P., DeRubeis, R. J., and Baraldi, A. N. (2016). Mediation analysis with survival outcomes: accelerated failure time vs. proportional hazards models. *Frontiers in Psychology*, 7, [423].

Gerestein, C. G., Eijkemans, M. J., de Jong, D., van der Burg, M. E., Dykgraaf, R. H., Kooi, G. S., Baalbergen, A., Burger, C. W., and Ansink, A. C. (2009). The prediction of progression-free and overall survival in women with an advanced stage of epithelial ovarian carcinoma. *British Journal of Obstetrics and Gynaecology*, *116*(3), 372–380.

Godskesen, T., Hansson, M. G., Nygren, P., Nordin, K., and Kihlbom, U. (2015). Hope for a cure and altruism are the main motives behind participation in phase 3 clinical cancer trials. *European Journal of Cancer Care*, *24*(1), 133–141.

Goff, B. A., Matthews, B. J., Larson, E. H., Andrilla, C. H., Wynn, M., Lishner, D. M., and Baldwin, L. M. (2007). Predictors of comprehensive surgical treatment in patients with ovarian cancer. *Cancer*, *109*(10), 2031–2042.

Griffiths, C. T. (1975). Surgical resection of tumor bulk in the primary treatment of ovarian carcinoma. *National Cancer Institute Monograph*, *42*, 101–104.

Griffiths, C. T., and Fuller, A. F. (1978). Intensive surgical and chemotherapeutic management of advanced ovarian cancer. *The Surgical clinics of North America*, *58*(1), 131–142.

Grimshaw, J. M., and Thomson, M. A. (1998). What have new efforts to change professional practice achieved? *Journal of the Royal Society of Medicine*, *91*(Suppl 35), 20–25.

Grimshaw, J. M., Eccles, M. P., Lavis, J. N., Hill, S. J., and Squires, J. E. (2012). Knowledge translation of research findings. *Implementation Science*, *7*, [50].

Grol, R. (2001). Successes and failures in the implementation of evidence-based guidelines for clinical practice. *Medical Care*, *39*(8), II46–II54.

Grol, R., and Grimshaw, J. (2003). From best evidence to best practice: effective implementation of change in patients' care. *The Lancet, 362*(9391), 1225–1230.

Gruen, R. L., Pitt, V., Green, S., Parkhill, A., Campbell, D., and Jolley, D. (2009). The effect of provider case volume on cancer mortality: systematic review and meta-analysis. *CA: A Cancer Journal for Clinicians*, *59*(3), 192–211.

Gupta, D., Lis, C. G., Vashi, P. G., and Lammersfeld, C. A. (2010). Impact of improved nutritional status on survival in ovarian cancer. *Supportive Care in Cancer*, *18*(3), 373–381.

Hafeman, D. M., and Schwartz, S. (2009). Opening the Black Box: a motivation for the assessment of mediation. *International Journal of Epidemiology*, *38*(3), 838–845.

Hanney, S., Boaz, A., Soper, B., and Jones, T. (2013). Engagement in research: an innovative three-stage review of the benefits for health-care performance. *Health Services and Delivery Research, 1,* [8].

Harding, K., Lynch, L., Porter, J., and Taylor, N. F. (2017). Organisational benefits of a strong research culture in a health service: a systematic review. *Australian Health Review*, *41*(1), 45–53.

Harter, P., Hilpert, F., Pfisterer, J., Sehouli, J., Lamparter, C., Kerkmann, M., and du Bois A. (2020). Therapiequalität des fortgeschrittenen Ovarialkarzinoms in Deutschland. *Frauenarzt*, *61*(3), 182–188.

Heaton, J., Day, J., and Britten, N. (2015). Inside the "black box" of a knowledge translation program in applied health research. *Qualitative Health Research*, *25*(11), 1477–1491.

Heintz, A. P., Odicino, F., Maisonneuve, P., Quinn, M. A., Benedet, J. L., Creasman, W. T., Ngan, H. Y., Pecorelli, S., and Beller, U. (2006). Carcinoma of the ovary. FIGO 26th Annual Report on the Results of Treatment in Gynecological Cancer. *International Journal of Gynaecology and Obstetrics*, *95*(Suppl 1), S161–S192.

Herzog, T. J., Armstrong, D. K., Brady, M. F., Coleman, R. L., Einstein, M. H., Monk, B. J., Mannel, R. S., Thigpen, J. T., Umpierre, S. A., Villella, J. A., and Alvarez, R. D. (2014). Ovarian cancer clinical trial endpoints: Society of Gynecologic Oncology white paper. *Gynecologic Oncology*, *132*(1), 8–17.

Hillner, B. E., Smith, T. J., and Desch, C. E. (2000). Hospital and physician volume or specialization and outcomes in cancer treatment: importance in quality of cancer care. *Journal of Clinical Oncology*, *18*(11), 2327–2340.

Højsgaard, S., Halekoh, U., and Yan, J. (2006). The R package geepack for generalized estimating equations. *Journal of Statistical Software, 15*(2), 1–11.

Hoskins, W. J., Bundy, B. N., Thigpen, J. T., and Omura, G. A. (1992). The influence of cytoreductive surgery on recurrence-free interval and survival in small-volume stage III epithelial ovarian cancer: a Gynecologic Oncology Group study. *Gynecologic Oncology, 47*(2), 159–166.

Iacobucci, D. (2008). *Mediation Analysis*. Thousand Oaks: Sage.

IBM Corp (2010). *IBM SPSS Statistics for Windows, Version 19.0*. Armonk, NY: IBM Corp.

Imai, K., Keele, L., and Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods, 15*(4), 309–334.

International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use. (2023). *Guideline for Good Clinical Practice (GCP) E6(R3).*
URL: https://database.ich.org/sites/default/files/ICH_E6%28R3%29_DraftGuideline_2023_0519.pdf
[accessed 30 August 2023].

Janni, W., Kiechle, M., Sommer, H., Rack, B., Gauger, K., Heinrigs, M., Steinfeld, D., Augustin, D., Simon, W., Harbeck, N., Friese, K., and ADEBAR Study Group (2006). Study participation improves treatment strategies and individual patient care in participating centers. *Anticancer Research*, *26*(5B), 3661–3667.

Jasen, P. (2009). From the "silent killer" to the "whispering disease": ovarian cancer and the uses of metaphor. *Medical History, 53*(4), 489–512.

Jha, P., Deboer, D., Sykora, K., and Naylor, C. D. (1996). Characteristics and mortality outcomes of thrombolysis trial participants and nonparticipants: a population-based comparison. *Journal of the American College of Cardiology*, *27*(6), 1335–1342.

Judd, C. M., and Kenny, D. A. (1981). Process analysis—estimating mediation in treatment evaluations. *Evaluation Review, 5*(5), 602–619.

Kaku, T., Ogawa, S., Kawano, Y., Ohishi, Y., Kobayashi, H., Hirakawa, T., and Nakano, H. (2003). Histological classification of ovarian cancer. *Medical Electron Microscopy*, *36*(1), 9–17.

Kaplan, E. L., and Meier, P. (1958) Nonparametric Estimation from incomplete observations. *Journal of the American Statistical Association*, *53*(282), 457–481.

Karam, A., Ledermann, J. A., Kim, J. W., Sehouli, J., Lu, K., Gourley, C., Katsumata, N., Burger, R. A., Nam, B. H., Bacon, M., Ng, C., Pfisterer, J., Bekkers, R., Casado Herráez, A., Redondo, A., Fujiwara, H., Gleeson, N., Rosengarten, O., Scambia, G., Zhu, J., Okamoto, A., Stuart, G., and Ochiai, K. on behalf of the participants of the 5th Ovarian Cancer Consensus Conference (2017). Fifth Ovarian Cancer Consensus Conference of the Gynecologic Cancer InterGroup: first-line interventions. *Annals of Oncology*, *28*(4), 711–717.

Karjalainen, S., and Palva, I. (1989). Do treatment protocols improve end results? A study of survival of patients with multiple myeloma in Finland. *BMJ*, *299*(6707), 1069–1072.

Katsumata, N., Yasuda, M., Takahashi, F., Isonishi, S., Jobo, T., Aoki, D., Tsuda, H., Sugiyama, T., Kodama, S., Kimura, E., Ochiai, K., Noda, K., and Japanese Gynecologic Oncology Group (2009). Dose-dense paclitaxel once a week in combination with carboplatin every 3 weeks for advanced ovarian cancer: a phase 3, open-label, randomised controlled trial. *The Lancet*, *374*(9698), 1331–1338.

Kaufman, J. S., Maclehose, R. F., and Kaufman, S. (2004). A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. *Epidemiologic Perspectives & Innovations*, *1,* [4].

Kelemen, L. E., and Köbel, M. (2011). Mucinous carcinomas of the ovary and colorectum: different organ, same dilemma. *The Lancet Oncology*, *12*(11), 1071–1080.

Kipps, E., Tan, D. S., and Kaye, S. B. (2013). Meeting the challenge of ascites in ovarian cancer: new avenues for therapy and research. *Nature Reviews Cancer*, *13*(4), 273–282.

Kistin, C., and Silverstein, M. (2015). Pilot studies: a critical but potentially misused component of interventional research. *Journal of the American Medical Association*, *314*(15), 1561–1562.

Krzyzanowska, M. K., Kaplan, R., and Sullivan, R. (2011). How may clinical research improve health care outcomes? *Annals of Oncology*, *22*(Suppl 7), vii10–vii15.

Landon, B. E., Wilson, I. B., and Cleary, P. D. (1998). A conceptual model of the effects of health care organizations on the quality of medical care. *Journal of the American Medical Association*, *279*(17), 1377–1382.

Landon, B. E., Zaslavsky, A. M., Beaulieu, N. D., Shaul, J. A., and Cleary, P. D. (2001). Health plan characteristics and consumers' assessments of quality. *Health Affairs*, *20*(2), 274–286.

Lange, T., Vansteelandt, S., and Bekaert, M. (2012). A simple unified approach for estimating natural direct and indirect effects. *American Journal of Epidemiology*, *176*(3), 190–195.

Lange, T., Rasmussen, M., and Thygesen, L. C. (2014). Assessing natural direct and indirect effects through multiple pathways. *American Journal of Epidemiology*, *179*(4), 513–518.

Lapointe-Shaw, L., Bouck, Z., Howell, N. A., Lange, T., Orchanian-Cheff, A., Austin, P. C., Ivers, N. M., Redelmeier, D. A., and Bell, C. M. (2018). Mediation analysis with a time-to-event outcome: a review of use and reporting in healthcare research. *BMC Medical Research Methodology, 18,* [118].

Lawler, M., Davies, L., Oberst, S., Oliver, K., Eggermont, A., Schmutz, A., La Vecchia, C., Allemani, Lievens, Y., Naredi, P., Cufer, T., Aggarwal, A., Aapro, M., Apostolidis, K., Baird, A.-M., Cardoso, F., Charalambous, A., Coleman, M. P., Costa, A., Crul, M., Dégi, C. L., Di Nicolantonio, F. , Erdem, S., Geanta, M., Geissler, J., Jassem, J., Jagielska, B., Jonsson, B., Kelly, D., Kelm, O., Kolarova, T., Kutluk, T., Lewison, G., Meunier, F., Pelouchova, J., Philip, T., Price, R., Rau, B., Rubio, I. T., Selby, P., Južnič Sotlar, M., Spurrier-Bernard, G., Van Hoeve, J. C., Vrdoljak, E., Westerhuis, W., Wojciechowska, U., and Sullivan, R. (2023). European Groundshot—addressing Europe's cancer research challenges: a Lancet Oncology Commission. *The Lancet Oncology, 24*(1), e11–e56.

Lawson, E. F., and Yazdany, J. (2012). Health care quality in systemic lupus erythematosus: using Donabedian's conceptual framework to understand what we know. *International Journal of Clinical Rheumatology*, *7*(1), 95–107.

Ledermann, J. A., Raja, F. A., Fotopoulou, C., Gonzalez-Martin, A., Colombo, N., Sessa, C., and ESMO Guidelines Working Group (2013). Newly diagnosed and relapsed epithelial ovarian carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*, *24*(Suppl 6), vi24–vi32.

Lee, M., Kim, S. W., Nam, E. J., Cho, H., Kim, J. H., Kim, Y. T., and Kim, S. (2014). ATP-based chemotherapy response assay in primary or recurrent ovarian and peritoneal cancer. *Yonsei Medical Journal*, *55*(6), 1664–1671.

Lee, H., Cashin, A. G., Lamb, S. E., Hopewell, S., Vansteelandt, S., VanderWeele, T. J., MacKinnon, D. P., Mansell, G., Collins, G., Golub, R. M., McAuley, J. H., and the AGReMA group (2021). A guideline for reporting mediation analyses of randomized trials and observational studies: the AGReMA statement. Journal of the American Medical Association, 326(11), 1045–1056.

Lipsky, A. M., and Greenland, S. (2022). Causal directed acyclic graphs. *Journal of the American Medical Association, 327*(11), 1083–1084.

Lohr, K. N. (1990). *Medicare: A Strategy for Quality Assurance, Volume I*. Washington, DC: National Academy Press.

Lund, B., Williamson, P., van Houwelingen, H. C., and Neijt, J. P. (1990). Comparison of the predictive power of different prognostic indices for overall survival in patients with advanced ovarian carcinoma. *Cancer Research*, *50*(15), 4626–4629.

Mackay, H. J., Brady, M. F., Oza, A. M., Reuss, A., Pujade-Lauraine, E., Swart, A. M., Siddiqui, N., Colombo, N., Bookman, M. A., Pfisterer, J., and du Bois, A. (2010). Prognostic relevance of uncommon ovarian histology in women with stage III/IV epithelial ovarian cancer. *International Journal of Gynecological Cancer*, *20*(6), 945–952.

MacKinnon, D. P., Fairchild, A. J., and Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., and Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*(1), 83–104.

Mahner, S., Harter, P., du Bois, A., Hilpert, F., Kerkmann, M., Sehouli, J., de Gregorio, N., Hanker, L., Heitz, F., Marmé, F., Woelber, L., Holtmann, L., Elser, G., Pfisterer, J., and AGO Study Group. Treatment and outcome of patients with high-grade advanced ovarian cancer (AOC): Real-world data in Germany (QS OVAR of the AGO Study Group) (2022). *Journal of Clinical Oncology*, *40*(Suppl 16), e17613–e17613.

Majumdar, S. R., Chang, W. C., and Armstrong, P. W. (2002). Do the investigative sites that take part in a positive clinical trial translate that evidence into practice? *The American Journal of Medicine*, *113*(2), 140–145.

Majumdar, S. R., Roe, M. T., Peterson, E. D., Chen, A. Y., Gibler, W. B., and Armstrong, P. W. (2008). Better outcomes for patients treated at hospitals that participate in clinical trials. *Archives of Internal Medicine*, *168*(6), 657–662.

Malvezzi, M., Carioli, G., Rodriguez, T., Negri, E., and La Vecchia, C. (2016). Global trends and predictions in ovarian cancer mortality. *Annals of Oncology, 27*(11), 2017–2025.

Manzanera, R., Moya, D., Guilabert, M., Plana, M., Gálvez, G., Ortner, J., and Mira, J. J. (2018). Quality Assurance and Patient Safety Measures: A Comparative Longitudinal Analysis. *International Journal of Environmental Research and Public Health*, *15*(8), [1568].

Marso, S. P., Daniels, G. H., Brown-Frandsen, K., Kristensen, P., Mann, J. F., Nauck, M. A., Nissen, S. E., Pocock, S., Poulter, N. R., Ravn, L. S., Steinberg, W. M., Stockner, M., Zinman, B., Bergenstal, R. M., and Buse, J. B. (2016). Liraglutide and cardiovascular outcomes in type 2 diabetes. *New England Journal of Medicine, 375*(4), 311–322.

Maxwell, S. E., and Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*(1), 23–44.

Meyer, J. I., Kennedy, A. W., Friedman, R., Ayoub, A., and Zepp, R. C. (1995). Ovarian carcinoma: value of CT in predicting success of debulking surgery. *American Journal of Roentgenology*, *165*(4), 875–878.

McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., and Fisher, P. (2007). The Hawthorne Effect: a randomised, controlled trial. *BMC Medical Research Methodology*, *7*, [30].

McGowan, L., Lesher, L. P., Norris, H. J., and Barnett, M. (1985). Misstaging of ovarian cancer. *Obstetrics and Gynecology*, *65*(4), 568–572.

McKibbon, K. A., Lokker, C., Wilczynski, N. L., Ciliska, D., Dobbins, M., Davis, D. A., Haynes, R. B., and Straus, S. E. (2010). A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006: a Tower of Babel? *Implementation Science, 5*, [16].

Mitchell, P. H., Ferketich, S., and Jennings, B. M. (1998). Quality health outcomes model. *Image–The Journal of Nursing Scholarship*, *30*(1), 43–46.

Moher, D., Glasziou, P., Chalmers, I., Nasser, M., Bossuyt, P., Korevaar, D. A., Graham, I. D., Ravaud, P., and Boutron, I. (2016). Increasing value and reducing waste in biomedical research: who's listening? *The Lancet*, *387*(10027), 1573–1586.

Morris, Z. S., Wooding, S., and Grant, J. (2011). The answer is 17 years, what is the question: understanding time lags in translational research. *Journal of the Royal Society of Medicine*, *104*(12), 510–520.

National Cancer Institute (2016). *Taking Part in Cancer Treatment Research Studies* (NIH Publication No. 16-6249). URL: https://www.cancer.gov/publications/patient-education/crs.pdf [last accessed 30 August 2023].

National Cancer Institute (2023). Surveillance, Epidemiology, and End Results Program. URL: https://seer.cancer.gov/ [last accessed 30 August 2023].

Nguyen, T. Q., Schmid, I., and Stuart, E. A. (2021). Clarifying causal mediation analysis for the applied researcher: Defining effects based on what we want to learn. *Psychological Methods, 26*(2), 255–271.

NIH Consensus Conference (1995). Ovarian cancer. Screening, treatment, and follow-up. NIH Consensus Development Panel on Ovarian Cancer. *Journal of the American Medical Association*, *273*(6), 491–497.

Ocana, A., and Tannock, I. F. (2011). When are "positive" clinical trials in oncology truly positive? *Journal of the National Cancer Institute*, *103*(1), 16–20.

Ozols, R. F., Bundy, B. N., Greer, B. E., Fowler, J. M., Clarke-Pearson, D., Burger, R. A., Mannel, R. S., DeGeest, K., Hartenbach, E. M., Baergen, R., and Gynecologic Oncology Group (2003). Phase III trial of carboplatin and paclitaxel compared with cisplatin and paclitaxel in patients with optimally resected stage III ovarian cancer: a Gynecologic Oncology Group study. *Journal of Clinical Oncology*, *21*(17), 3194–3200.

Ozols, R. F., Rubin, S., Thomas, G., and Robboy, S. (2006). Epithelial ovarian cancer. In Hoskins, W. J., Perez, C. A., Young, R. C., Barakat, R., Markman, M., and Randall M. (Eds), *Principles and Practice of Gynecologic Oncology* (4th ed). Philadelphia: Lippincott Williams & Wilkins, p. 895–987.

Pater, J., Rochon, J., Parmar, M. K., and Selby, P. (2011). Future research and methodological approaches. *Annals of Oncology*, *22*(Suppl 7), vii57–vii61.

Partridge, E. E., and Barnes, M. N. (1999). Epithelial ovarian cancer: prevention, diagnosis, and treatment. *CA: A Cancer Journal for Clinicians*, *49*(5), 297–320.

Pearl, J. (2012). The causal mediation formula—a guide to the assessment of pathways and mechanisms. *Prevention Science, 13*(4), 426–36.

Pecorelli, S., Benedet, J. L., Creasman, W. T., and Shepherd, J. H. on behalf of the 1994–1997 FIGO Committee on Gynecologic Oncology (1999). FIGO staging of gynecologic cancer. *International Journal of Gynecology & Obstetrics, 65*(3), 243–249.

Peppercorn, J. M., Weeks, J. C., Cook, E. F., and Joffe, S. (2004). Comparison of outcomes in cancer patients treated within and outside clinical trials: conceptual framework and structured review. *The Lancet*, *363*(9405), 263–270.

Pfisterer, J., Weber, B., Reuss, A., Kimmig, R., du Bois, A., Wagner, U., Bourgeois, H., Meier, W., Costa, S., Blohmer, J. U., Lortholary, A., Olbricht, S., Stähle, A., Jackisch, C., Hardy-Bessard, A. C., Möbus, V., Quaas, J., Richter, B., Schröder, W., Geay, J. F., Lueck, H. J., Kuhn, W., Meden, H., Nitz, U., and Pujade-Lauraine, E. (2006). Randomized phase III trial of topotecan following carboplatin and paclitaxel in first-line treatment of advanced ovarian cancer: a gynecologic cancer intergroup trial of the AGO-OVAR and GINECO. *Journal of the National Cancer Institute*, *98*(15), 1036–1045.

Piccart, M. J., Bertelsen, K., James, K., Cassidy, J., Mangioni, C., Simonsen, E., Stuart, G., Kaye, S., Vergote, I., Blom, R., Grimshaw, R., Atkinson, R. J., Swenerton, K. D., Trope, C., Nardi, M., Kaern, J., Tumolo, S., Timmers, P., Roy, J. A., Lhoas, F., Lindvall, B., Bacon, M., Birt, A., Andersen, J. E., Zee, B., Paul, J., Baron, B., and Pecorelli, S. (2000). Randomized intergroup trial of cisplatin-paclitaxel versus cisplatin-cyclophosphamide in women with advanced epithelial ovarian cancer: three-year results. *Journal of the National Cancer Institute*, *92*(9), 699–708.

Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology, 66,* 825–852.

Prentice, R. L. (1989). Surrogate endpoints in clinical trials: definition and operational criteria. *Statistics in Medicine*, *8*(4), 431–440.

Puls, L. E., Duniho, T., Hunter, J. E., Kryscio, R., Blackhurst, D., and Gallion, H. (1996). The prognostic implication of ascites in advanced-stage ovarian cancer. *Gynecologic Oncology*, *61*(1), 109–112.

Qin, X. (in press). Sample size and power calculations for causal mediation analysis: A Tutorial and Shiny App. *Behavior Research Methods*.

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Reade, C. J. and Elit, L. M. (2012). Trends in gynecologic cancer care in North America. *Obstetrics and Gynecology Clinics of North America, 39*(2), 107–129.

Reade, C. J. and Elit, L. M. (2019). Current Quality of Gynecologic Cancer Care in North America. *Obstetrics and Gynecology Clinics of North America,* 4*6*(1), 1–17.

Rijnhart, J. J., Lamp, S. J., Valente, M. J., MacKinnon, D. P., Twisk, J. W., and Heymans, M. W. (2021). Mediation analysis methods used in observational research: a scoping review and recommendations. *BMC Medical Research Methodology, 21,* [226].

Robins, J. M., and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, *3*(2), 143–155.

Rochon, J., and du Bois, A. (2011). Clinical research in epithelial ovarian cancer and patients' outcome. *Annals of Oncology*, *22*(Suppl 7), vii16–vii19.

Rochon, J., du Bois, A., and Lange, T. (2014). Mediation analysis of the relationship between institutional research activity and patient survival. *BMC Medical Research Methodology*, *14*, [9].

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomised studies. *Journal of Educational Psychology*, *66*(55), 688–701.

Rubin, D. B. (2004). Direct and indirect causal effects via potential outcomes. *Scandinavian Journal of Statistics: Theory and Applications*, *31*(2), 161–170.

Ruesch, J. (1961). Psychosomatic medicine and the behavioral sciences. *Psychosomatic Medicine*, *23*(4), 277–288.

Saltz, L. B. (2008). Progress in cancer care: the hope, the hype, and the gap between reality and perception. *Journal of Clinical Oncology, 26*(31), 5020–5021.

Schemper, M., and Smith, T. L. (1996). A note on quantifying follow-up in studies of failure time. *Controlled clinical trials*, *17*(4), 343–346.

Sehouli, J., du Bois, A., Pfisterer, J., Hilpert, F., Kerkmann, M., Mahner, S., de Gregorio, N., Hanker, L. C., Heitz, F., Marmé, F., Woelber, L. L., Holtmann, L., Polleis, S., Wimberger, P., and Harter, P. (2022) Real world data of treatment and outcome of patients with early ovarian cancer (FIGO I) in Germany (QS Ovar of the AGO Study Group). *International Journal of Gynecologic Cancer, 32,* [A228].

Selby, P. (2011). The impact of the process of clinical research on health service outcomes. *Annals of Oncology*, *22*(Suppl 7), vii2–vii4.

Selby, P., and Autier, P. (2011). The impact of the process of clinical research on health service outcomes. *Annals of Oncology*, *22*(Suppl 7), vii5–vii9.

Selby, P., Liu, L., Downing, A., Banks, I., Wilson, R., Stephens, R., Meunier, F., Rochon, J., Morris, E., Seymour, M., Gregory, W., Lawler, M., and Boaz, A. (2019a). How can clinical research improve European health outcomes in cancer? *Journal of Cancer Policy*, *20*, [100182].

Selby, P., Popescu, R., Lawler, M., Butcher, H., and Costa, A. (2019b). The value and future developments of multidisciplinary team cancer care. *American Society of Clinical Oncology Educational Book, 39,* 332–340.

Shen-Gunther, J., and Mannel, R. S. (2002). Ascites as a predictor of ovarian malignancy. *Gynecologic Oncology*, *87*(1), 77–83.

Shimizu, Y., Kamoi, S., Amada, S., Akiyama, F., and Silverberg, S. G. (1998). Toward the development of a universal grading system for ovarian epithelial carcinoma: testing of a proposed system in a series of 461 patients with uniform treatment and follow-up. *Cancer*, *82*(5), 893–901.

Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., and Moore, L. (2021). A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *The BMJ*, *374*, [n2061].

Sørensen, J. B., Klee, M., Palshof, T., and Hansen, H. H. (1993). Performance status assessment in cancer patients. An inter-observer variability study. *British Journal of Cancer*, *67*(4), 773–775.

Splawa-Neyman, J., Dabrowska, D. M., and Speed, T. P. (1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statistical Science*, *5*(4), 465–472.

Staebler, A., and Mayr, D. (2017). [The 2016 update of the S3 guideline for malignant tumours of the ovary: Role of pathology in diagnosis, therapy and clinical management of epithelial tumours]. *Der Pathologe*, *38*(4), 331–344.

Stiller, C. A. (1994). Centralised treatment, entry to trials and survival. *British Journal of Cancer*, *70*(2), 352–362.

Stuart, E. A., Schmid, I., Nguyen, T., Sarker, E., Pittman, A., Benke, K., Rudolph, K., Badillo-Goicoechea, E., and Leoutsakos, J. M. (2021). Assumptions not often assessed or satisfied in published mediation analyses in psychology and psychiatry. *Epidemiologic Reviews, 43(1),* 48–52.

Sung, N. S., Crowley Jr, W. F., Genel, M., Salber, P., Sandy, L., Sherwood, L. M., Johnson, S. B., Catanese, V., Tilson, H., Getz, K., Larson, E. L., Scheinberg, D., Reece, E. A., Slavkin, H., Dobs, A., Grebb, J., Martinez, R. A., Korn, A., and Rimoin, D. (2003). Central challenges facing the national clinical research enterprise. *Journal of the American Medical Association*, *289*(10), 1278–1287.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 71*(3), 209–249.

Teramukai, S., Ochiai, K., Tada, H., Fukushima, M., and Japan Multinational Trial Organization OC01-01 (2007). PIEPOC: a new prognostic index for advanced epithelial ovarian cancer–Japan Multinational Trial Organization OC01-01. *Journal of Clinical Oncology*, *25*(22), 3302–3306.

Therneau, T. (2013). *A Package for Survival Analysis in S*. R package version 2.37-4.

Thigpen, T., Brady, M. F., Omura, G. A., Creasman, W. T., McGuire, W. P., Hoskins, W. J., and Williams, S. (1993). Age as a prognostic factor in ovarian carcinoma. The Gynecologic Oncology Group experience. *Cancer*, *71*(Suppl 2), 606–614.

Truong, T. H., Weeks, J. C., Cook, E. F., and Joffe, S. (2011). Altruism among participants in cancer clinical trials. *Clinical Trials*, *8*(5), 616–623.

Van de Laar, R., IntHout, J., Van Gorp, T., Verdonschot, S., van Altena, A. M., Gerestein, C. G., Massuger, L. F., Zusterzeel, P. L., and Kruitwagen, R. F. (2014). External validation of three prognostic models for overall survival in patients with advanced-stage epithelial ovarian cancer. *British Journal of Cancer*, *110*(1), 42–48.

VanderWeele, T. J., and Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface, 2*(4):457–68.

Vanderweele, T. J., and Vansteelandt, S. (2010). Odds ratios for mediation analysis for a dichotomous outcome. *American Journal of Epidemiology, 172*(12), 1339–48.

VanderWeele, T. J. (2011). Causal mediation analysis with survival data. *Epidemiology*, *22*(4), 582–585.

VanderWeele, T. J. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York, NY: Oxford University Press.

VanderWeele, T. J. (2016). Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health, 37*, 17–32.

Vansteelandt, S., Bekaert, M., and Lange, T. (2012). Imputation strategies for the estimation of natural direct and indirect effects. *Epidemiologic Methods, 1*(1), 131–58.

Vansteelandt, S., Linder, M., Vandenberghe, S., Steen, J., and Madsen, J. (2019). Mediation analysis of time-to-event endpoints accounting for repeatedly measured mediators subject to time-varying confounding. *Statistics in Medicine, 38*(24), 4828–4840.

Vergote, I., Gonzalez-Martin, A., Lorusso, D., Gourley, C., Mirza, M. R., Kurtz, J. E., Okamoto, A., Moore, K., Kridelka, F., McNeish, I., Reuss, A., Votan, B., du Bois, A., Mahner, S., Ray-Coquard, I., Kohn, E. C.,

Berek, J. S., Tan, D. S. P., Colombo, N., Zang, R., Concin, N., O'Donnell, D., Rauh-Hain, A., Herrington, C. S., Marth, C., Poveda, A., Fujiwara, K., Stuart, G. C. E., Oza, A. M., Bookman, M. A., and Zhu, J. (2022). Clinical research in ovarian cancer: consensus recommendations from the Gynecologic Cancer Inter-Group. *The Lancet Oncology, 23*(8), e374-e384.

Vist, G. E., Bryant, D., Somerville, L., Birminghem, T., and Oxman, A. D. (2008). Outcomes of patients who participate in randomized controlled trials compared to similar patients receiving similar interventions who do not participate. *The Cochrane Database of Systematic Reviews*, (3), [MR000009].

Vo, T. T., Superchi, C., Boutron, I., and Vansteelandt, S. (2020). The conduct and reporting of mediation analysis in recently published randomized controlled trials: results from a methodological systematic review. *Journal of Clinical Epidemiology, 117,* 78–88.

Wagner, U., Harter, P., Hilpert, F., Mahner, S., Reuss, A., du Bois, A., Petru, E., Meier, W., Ortner, P., König, K., Lindel, K., Grab, D., Piso, P., Ortmann, O., Runnebaum, I., Pfisterer, J., Lüftner, D., Frickhofen, N., Grünwald, F., Maier, B. O., Diebold, J., Hauptmann, S., Kommoss, F., Emons, G., Radeleff, B., Gebhardt, M., Arnold, N., Calaminus, G., Weisse, I., Weis, J., Sehouli, J., Fink, D., Burges, A., Hasenburg, A., and Eggert, C. (2013). S3-Guideline on Diagnostics, Therapy and Follow-up of Malignant Ovarian Tumours: Short version 1.0 - AWMF registration number: 032/035OL, June 2013. *Geburtshilfe und Frauenheilkunde*, *73*(9), 874–889.

Ward, L. C., Fielding, J. W., Dunn, J. A., and Kelly, K. A. (1992). The selection of cases for randomised trials: a registry survey of concurrent trial and non-trial patients. The British Stomach Cancer Group. *British Journal of Cancer*, *66*(5), 943–950.

Winter III, W. E., Maxwell, G. L., Tian, C., Carlson, J. W., Ozols, R. F., Rose, P. G., Markman, M., Armstrong, D. K., Muggia, F., McGuire, W. P., and Gynecologic Oncology Group Study (2007). Prognostic factors for stage III epithelial ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology*, *25*(24), 3621–3627.

Woolf, S. H. (2008). The meaning of translational research and why it matters. *Journal of the American Medical Association*, *299*(2), 211–213.

Wright, S. (1934). The method of path coefficients. *The Annals of Mathematical Statistics, 5*(3), 161–215.

Yancik, R. (1993). Ovarian cancer. Age contrasts in incidence, histology, disease stage at diagnosis, and mortality. *Cancer*, *71*(Suppl 2), 517–523.

Zhao, X., Lynch Jr, J. G., and Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197–206.

Zon, R., Meropol, N. J., Catalano, R. B., and Schilsky, R. L. (2008). American Society of Clinical Oncology Statement on minimum standards and exemplary attributes of clinical trial sites. *Journal of Clinical Oncology*, *26*(15), 2562–2567.

7 Own publications

This dissertation thesis was done in the context of a German quality assurance program ("Quality Assurance of Ovarian Cancer Treatment", QS-OVAR). QS-OVAR is an initiative of the Arbeitsgemeinschaft Gynaekologische Onkologie (AGO) Organkommission OVAR, a subcommittee of the German Cancer Society. The aim of this program is to describe the pattern and quality of care of patients with ovarian cancer in Germany as well as to improve their outcomes.

I served as principal investigator and acted as responsible biostatistician for QS-OVAR 2001, 2004, and 2008. I conducted all the pre-work including an extensive literature review, wrote the study protocol and submitted it to the Ethics Committee of the Medical Faculty of the University of Heidelberg for review. The Ethics Committee had no objections against the proposal and approved the study (reference number: S-446/2013). I contributed to the concept and study design as well as to the design of the data collection forms for QS-OVAR. Medizinische Marktforschung (MMF GmbH, Dortmund) was responsible for data management. Data were collected on paper report forms, cross-checked with surgical and pathological reports, and then entered into an electronic database. Trained data managers verified the accuracy and plausibility of the data. The use of QS-OVAR data for my dissertation was approved by Prof. Andreas du Bois on behalf of AGO-OVAR. In collaboration with MMF GmbH, I carried out further plausibility and completeness analyses of the individual data sets, regularly updated the treatment definitions according to respective guidelines, maintained follow-up data in existing data sets, merged the databases, and prepared the data for the main analysis as well as for the sensitivity analyses. I was responsible for the planning and programming of all statistical analyses, conduct of the analyses, the presentation and interpretation of the results for the individual QS-OVAR cohorts as well as the pooled data. Finally, I drafted the two primary publications with me as first author and contributed to several secondary publications (see details below).

During my dissertation project, I worked on medical and health care topics as well as on methodological aspects of QS-OVAR. Together with Prof. Andreas du Bois and the AGO-OVAR, I worked on clinical studies in gynecological cancers and questions related to diagnosis, patterns and quality of care, and prognosis of ovarian cancer. During my employment as research assistant at the Institute for Medical Biometry in Heidelberg from April 2009 until September 2013, I worked on methodological aspects of quality of health care under the supervision of Prof. Meinhard Kieser. In addition, in 2012 I started a collaboration with Prof. Theis Lange (University of Copenhagen) to learn more about causal inference, to further develop the methodology, and eventually apply causal mediation analysis to the research question of my dissertation.

Results of this dissertation thesis were partly published in the following two articles:

- Rochon, J., and du Bois, A. (2011). Clinical research in epithelial ovarian cancer and patients' outcome. *Annals of Oncology*, *22*(Suppl 7), vii16–vii19.

  This publication was based on the first QS-OVAR cohort with complete 3-year follow-up data and focused on the outcome of ovarian cancer patients treated in research-active hospitals. My contribution to this publication was the design of the study, preparation, programming and conduct of all statistical analyses, the description and interpretation of the results, and the draft of the manuscript.

- Rochon, J., du Bois, A., and Lange, T. (2014). Mediation analysis of the relationship between institutional research activity and patient survival. *BMC Medical Research Methodology*, *14*, [9].

  This publication serves as the main output of my thesis, in which a novel approach is introduced for assessing causal mediation across various types of outcome variables, with a focus on time-to-event endpoints in the presence of two potential mediators. The analysis was conducted using data from 352 advanced ovarian cancer patients from the first QS-OVAR cohort. My contributions to this publication included the design of the study, data analysis, and drafting of the manuscript. The accompanying supplementary online material provides the R code and a comprehensive description of the analysis process, enabling other researchers to evaluate causal mediation in a survival setting. Last but not least, this material offers sensitivity analyses, shedding light on testing for interactions between exposure and confounders, examining the linearity assumption, and accounting for potential misspecifications in the mediators.

I recently co-authored two publications that were partially based on the insights gained through my work on the dissertation. My contribution to these two publications is focused on methodological aspects and on assessing the future challenges and opportunities of research in the respective area of interest. In addition, I helped to draft the manuscripts, contributed to their revisions, and approved the final versions.

- Burger, H. U., Gerlinger, C., Harbron, C., Koch, A., Posch, M., Rochon, J., and Schiel, A. (2021). The use of external controls: To what extent can it currently be recommended? *Pharmaceutical Statistics, 20*(6)*,* 1002–1016.
- Selby, P., Liu, L., Downing, A., Banks, I., Wilson, R., Stephens, R., Meunier, F., Rochon, J., Morris, E., Seymour, M., Gregory, W., Lawler, M., and Boaz, A. (2019). How can clinical research improve European health outcomes in cancer? *Journal of Cancer Policy, 20,* [100182].

My long collaboration with the AGO-OVAR is reflected by several international and German publications. As can be seen from the list below, I co-authored articles on clinical studies in gynecological cancers as well as articles about the quality of care in ovarian cancer using some data from QS-OVAR. Relevant results from these publications are cited in my dissertation thesis. In all these publications, my role was that of a responsible biostatistician: I was involved in planning and executing all statistical analyses, interpreting the results, assisting in drafting the manuscripts, contributing to the respective revisions, and approving the final versions. Lastly, I have made significant contributions to consensus statements on the management of ovarian cancer.

- du Bois, A., Rochon, J., Lamparter, C., Elser, G., and Pfisterer, J. (2009). Versorgungsstruktur, Qualitätssicherung und klinische Studien beim invasiven Ovarialkarzinom in Deutschland – Aktionsprogramm QS-OVAR der AGO und Aktivitäten der AGO Studiengruppe. In: R. Kreienberg, A. du Bois, J. Pfisterer, S. Schindelmann, and B. Schmalfeldt (Hrsg.) *Management des Ovarialkarzinoms.* Heidelberg: Springer Medizin Verlag, S. 325–338.
- du Bois, A., Rochon, J., Lamparter, C., and Pfisterer, J. (2009). Die Qualität der Therapie des Ovarialkarzinoms in Deutschland: Dritte Stufe der Qualitätssicherungserhebung QS-OVAR der Arbeitsgemeinschaft Gynäkologische Onkologie (AGO) Kommission OVAR. *Frauenarzt, 50*(9), 742–751.

- du Bois, A., Rochon, J., Lamparter, C., and Pfisterer, J. (2009). Die Qualität der Rezidivtherapie beim Ovarialkarzinom in Deutschland: Eine Qualitätssicherungserhebung im Rahmen des Aktionsprogramms der AGO Kommission OVAR (QS-OVAR). *Frauenarzt, 50*(11), 952–959.
- du Bois, A., Rochon, J., Lamparter, C., Pfisterer, J., and AGO Organkommission OVAR (2005). Pattern of care and impact of participation in clinical studies on the outcome in ovarian cancer. *International Journal of Gynecological Cancer, 15*(2), 183–191.
- du Bois, A., Rochon, J., Lamparter, C., and Pfisterer, J. (2005). Ovarialkarzinom – Versorgungsstruktur und -qualität in Deutschland 2001–2004. *Frauenarzt, 46*(7), 560–567.
- du Bois, A., Rochon, J., Lamparter, C., and Pfisterer, J. (2005). Das Qualitätssicherungsprogramm der Kommission OVAR der AGO (QS-OVAR): Versorgungsstruktur und Realität in Deutschland 2001. *Zentralblatt Gynäkologie, 127,* 9–17.
- du Bois, A., Rochon, J., Lamparter, C., and Pfisterer, J. (2005). Welchen Einfluss haben Klinikmerkmale auf die Prognose beim Ovarialkarzinom in Deutschland? *Zentralblatt für Gynäkologie, 127,* 18–30.
- du Bois, A., Rochon, J., Pfisterer, J., and Hoskins, W. J. (2009). Variations in institutional infrastructure, physician specialization and experience, and outcome in ovarian cancer: a systematic review. *Gynecologic Oncology, 112*(2), 422–436.
- du Bois, A., Quinn, M., Thigpen, T., Vermorken, J. B., Avall-Lundqvist, E., Bookman, M., Bowtell, D., Brady, M., Casado, A., Cervantes, A., Eisenhauer, E., Friedlaender, M., Fujiwara, K., Grenman, S., Guastalla, J. P., Harper, P., Hogberg, T., Kaye, S., Kitchener, H., Kristensen, G., Mannel, R., Meier, W., Miller, B., Neijt, J. P., Oza, A., Ozols, R., Parmar, M. K., Pecorelli, S., Pfisterer, J., Poveda, A., Provencher, D., Pujade-Lauraine, E., Randall, M., Rochon, J., Rustin, G., Sagae, S., Stehman, F., Stuart, G., Trimble, E., Vasey, P., Vergote, I., Verheijen, R., and Wagner, U. (2005). 2004 consensus statements on the management of ovarian cancer: final document of the 3rd International Gynecologic Cancer Intergroup Ovarian Cancer Consensus Conference (GCIG OCCC 2004). *Annals of Oncology, 16*(Suppl. 8), viii7–viii12.
- du Bois, A., Weber, B., Rochon, J., Meier, W., Gaupil, A., Olbricht, S., Berats, J. C., Kuhn, W., Orfeuvre, H., Wagner, U., Richter, B., Lueck, H. J., Pfisterer, J., Costa, S., Schroeder, W., Kimmig, R., and Pujade-Lauraine, E. for AGO-OVAR and GINECO (2006). Addition of epirubicin as a third drug to carboplatin-paclitaxel in first-line treatment of advanced ovarian cancer: A prospectively randomized Gynecologic Cancer Intergroup trial by the Arbeitsgemeinschaft Gynaekologische Onkologie Ovarian Cancer Study Group and the Groupe d'Investigateurs Nationaux pour l'Étude des Cancers Ovariens. *Journal of Clinical Oncology, 24*(7), 1127–1135.
- Heitz, F., du Bois, A., Rochon, J., Scheil-Bertram, S., Hils, R., Fisseler-Eckhoff, A., Barinoff, J., Kaub, C., and Harter, P. (2012). Requirements to assess feasibility of phase 0 trials during major abdominal surgery: variability of PARP activity. *Clinical Cancer Research, 18*(9), 2632–2637.
- Heitz, F., Rochon, J., Harter, P., Lueck, H. J., Fisseler-Eckhoff, A., Barinoff, J., Traut, A., Lorenz-Salehi, F., and du Bois, A. (2011). Cerebral metastases in metastatic breast cancer: disease-specific risk factors and survival. *Annals of Oncology, 22*(7), 1571–1581.
- Hilpert, F., Harter, P., Pfisterer, J., Rochon, J., Friedemann, J., Lamparter, C., and du Bois, A. (2010). Eigenverantwortliche Qualitätssicherung zur Therapiequalität des Ovarialkarzinoms in Deutschland. *Der Gynäkologe, 43*(12), 1022–1027.

- Hilpert, F., Harter, P., Pfisterer, J., Rochon, J., Friedemann, J., Lamparter, C., and du Bois, A. (2010). Qualitätssicherungserhebung QS-OVAR: Therapiestandards kritisch überdenken. *Im Focus Onkologie, 4,* 65–73.
- Kommoss, S., Rochon, J., Harter, P., Heitz, F., Grabowski, J. P., Ewald-Riegler, N., Haberstroh, M., Neunhoeffer, T., Barinoff, J., Gomez, R., Traut, A., and du Bois, A. (2010). Prognostic impact of extended surgical procedures in advanced-stage primary ovarian cancer. *Annals of Surgical Oncology, 17,* 279–286.
- Pater, J., Rochon, J., Parmar, M. K., and Selby, P. (2011). Future research and methodological approaches. *Annals of Oncology, 22*(Suppl 7), vii57–vii61.
- Pfisterer, J., Plante, M., Vergote, I., du Bois, A., Hirte, H., Lacave, A. J., Wagner, U., Stahle, A., Stuart, G., Kimmig, R., Olbricht, S., Le, T., Emerich, J., Kuhn, W., Bentley, J., Jackisch, C., Lueck, H. J., Rochon, J., Zimmermann, A. H., and Eisenhauer, E. (2006). Gemcitabine plus carboplatin compared with carboplatin in patients with platinum-sensitive recurrent ovarian cancer: an intergroup trial of the AGO-OVAR, the NCIC CTG, and the EORTC GCG. *Journal of Clinical Oncology, 24*(29), 4699–4707.
- Schade-Brittinger, C., Reuss, A., and Rochon, J. (2009). Studiendesign. In: Kreienberg, R., du Bois, A., Pfisterer, J., Schindelmann, S. und Schmalfeldt, B. (Hrsg.), *Management des Ovarialkarzinoms.* Heidelberg: Springer Medizin Verlag, S. 247–253.
- Stuart, G., Avall-Lundqvist, E., du Bois, A., Bookman, M., Bowtell, D., Brady, M., Casado, A., Cervantes, A., Eisenhauer, E., Friedlaender, M., Fujiwara, K., Grenman, S., Guastalla, J. P., Harper, P., Hogberg, T., Kaye, S., Kitchener, H., Kristensen, G., Mannel, R., Meier, W., Miller, B., Oza, A., Ozols, R., Parmar, M. K., Pfisterer, J., Poveda, A., Provencher, D., Pujade-Lauraine, E., Quinn, M., Randall, M., Rochon, J., Rustin, G., Sagae, S., Stehman, F., Trimble, E., Thigpen, T., Vasey, P., Vergote, I., Verheijen, R., Vermorken, J. B., and Wagner, U. (2005). 3rd International Ovarian Cancer Consensus Conference: outstanding issues for future consideration. *Annals of Oncology, 16(Suppl. 8),* viii36–viii38.
- Vermorken, J. B., Parmar, M. K., Brady, M. F., Eisenhauer, E. A., Hogberg, T., Ozols, R. F., Rochon, J., Rustin, G. J., Sagae, S., and Verheijen, R. H. (2005). Clinical trials in ovarian carcinoma: study methodology. *Annals of Oncology, 16(Suppl. 8),* viii20–viii29.

My additional individual and co-authored publications, encompassing a diverse range of topics and therapeutic areas, are listed below in alphabetical order. These publications not only showcase my work as a statistician in collaborative clinical trials and methodology groups, but also highlight my commitment to advancing medical research and improving patient outcomes through statistics.

- Eberbach, A., Becker, A., Rochon, J., Finkemeler, H., Wagner, A., and Donner-Banzhoff, N. (2016). A simple heuristic for internet-based evidence search in primary care: a randomized controlled trial. *Advances in Medical Education and Practice, 7,* 433–41.
- Erler, A., Beyer, M., Petersen, J. J., Saal, K., Rath, T., Rochon, J., Haefeli, W. E., and Gerlach, F. M. (2012). How to improve drug dosing for patients with renal impairment in primary care – A cluster-randomized controlled trial. *BMC Family Practice, 13,* [91].
- Feigl, G. C., Schebesch, K. M., Rochon, J., Warnat, J., Woertgen, C., Schmidt, S., Lange, M., Schlaier, J., and Brawanski, A. T. (2011). Analysis of risk factors influencing the development of severe dizziness in patients with vestibular schwannomas in the immediate postoperative phase. *Clinical Neurology and Neurosurgery, 113*(1), 52–56.

- Freund, T., Gondan, M., Rochon, J., Peters-Klimm, F., Campbell, S., Wensing, M., and Szecsenyi, J. (2013). Comparison of physician referral and insurance claims data-based risk prediction as approaches to identify patients for care management in primary care: an observational study. *BMC Family Practice, 14,* [157].
- Freund, T., Peters-Klimm, F., Boyd, C. M., Mahler, C., Gensichen, J., Erler, A., Beyer, M., Gondan, M., Rochon, J., Gerlach, F. M., and Szecsenyi, J. (2016). Medical assistant-based care management for high-risk patients in small primary care practices: a cluster randomized clinical trial. *Annals of Internal Medicine, 164*(5), 323–330.
- Freund, T., Peters-Klimm, F., Rochon, J., Mahler, C., Gensichen, J., Erler, A., Beyer, M., Baldauf, A., Gerlach, F. M., and Szecsenyi, J. (2011). Primary care practice-based care management for chronically ill patients (PraCMan): study protocol for a cluster randomized controlled trial [ISRCTN56104508]. *Trials, 12,* [163].
- Glockzin, G., Rochon, J., Arnold, D., Lang, S. A., Klebl, F., Zeman, F., Koller, M., Schlitt, H., and Piso, P. (2013). A prospective multicenter phase II study evaluating multimodality treatment of patients with peritoneal carcinomatosis arising from appendiceal and colorectal cancer: the COMBATAC trial. *BMC Cancer, 13,* [67].
- Goetz, K., Campbell, S. M., Willms, S., Rochon, J., Klingenberg, A., and Szecsenyi, J. (2011). How do chronically ill patients evaluate their medical care? An observational study with 46919 patients in 676 primary care practices. *International Journal of Person-Centered Medicine, 1,* 338–346.
- Haasenritter, J., Stanze, D., Widera, G., Wilimzig, C., Abu Hani, M., Sönnichsen, A., Bösner, S., Rochon, J., and Donner-Banzhoff, N. (2012). Does the patient with chest pain have a coronary heart disease? Diagnostic value of single symptoms and signs—a meta-analysis. *Croatian Medical Journal, 53,* 432–441.
- Hattemer, K., Knake, S., Reis, J., Rochon, J., Oertel, W., Rosenow, F., and Hamer, H. (2007). Excitability of the motor cortex during ovulatory and anovulatory cycles: a transcranial magnetic stimulation study. *Clinical Endocrinology, 66*(3), 387–393.
- Hoffmann, B., Müller, V., Rochon, J., Gondan, M., Müller, B., Albay, Z., Weppler, K., Leifermann, M., Mießner, C., Güthlin, C., Parker, D., Hofinger, G., and Gerlach, F. M. (2014). Effects of a team-based assessment and intervention on patient safety culture in general practice: an open randomised controlled trial. *BMJ Quality & Safety, 23*(1), 35–46.
- Keller, H., Krones, T., Becker, A., Hirsch, O., Sönnichsen, A. C, Popert, U., Kaufmann-Kolle, P., Rochon, J., Wegscheider, K., Baum, E., and Donner-Banzhoff, N. (2012). Arriba: effects of an educational intervention on prescribing behaviour in prevention of CVD in general practice. *European Journal of Preventive Cardiology, 19*(3), 322–329.
- Keller, H., Krones, T., Sönnichsen, A., Sadowski, E., Popert, U., Rochon, J., Kaufmann-Kolle, P., Szecsenyi, J., Baum, E., and Donner-Banzhoff, N. (2007). Medikamentöse Prävention von kardiovaskulären Erkrankungen: Verschreiben Hausärzte risikoangemessen? *Zeitschrift für Allgemeinmedizin, 83*(9), 359–364.
- Klafke, N., Mahler, C., von Hagens, C., Rochon, J., Schneeweiss, A., Müller, A., Salize, H. J., and Joos, S. (2015). A complex nursing intervention of complementary and alternative medicine (CAM) to increase quality of life in patients with breast and gynecologic cancer undergoing chemotherapy: study protocol for a partially randomized patient preference trial. *Trials, 16,* [51].

- Knake, S., Rochon, J., Fleischer, S., Katsarou, N., Back, T., Vescovi, M., Oertel, W., Reis, J., Hamer, H., and Rosenow, F. (2006). Status epilepticus after stroke is associated with increased long-term case fatality. *Epilepsia, 47*(12), 2020–2026.
- Krones, T., Keller, H., Sönnichsen, A., Sadowski, E., Baum, E., Wegscheider, K., Rochon, J., and Donner-Banzhoff, N. (2008). Absolute cardiovascular risk and shared decision-making in primary care: A single-blind cluster randomized controlled trial. *Annals of Family Medicine, 6,* 218–227.
- Ludt, S., Campbell, S. M., Petek, D., Rochon, J., Szecsenyi, J., van Lieshout, J., Wensing, M., and Ose, D. (2013). Which practice characteristics are associated with the quality of cardiovascular disease prevention in European primary care? *Implementation Science, 8*, [27].
- Ludt, S., van Lieshout, J., Campbell, S. M., Rochon, J., Ose, D., Freund, T., Wensing, M., and Szecsenyi, J. (2012). Identifying factors associated with experiences of coronary heart disease patients receiving structured chronic care and counselling in European primary care*. BMC Health Services Research, 12,* [221].
- Ludt, S., Wensing, M., Campbell, S. M., Ose, D., van Lieshout, J., Rochon, J., Uhlmann, L., and Szecsenyi, J. (2014). The challenge of cardiovascular prevention in primary care: implications of a European observational study in 8928 patients at different risk levels. *European Journal of Preventive Cardiology, 21*(2), 203–213.
- Ludt, S., Wensing, M., Szecsenyi, J., van Lieshout, J., Rochon, J., Freund, T., Campbell, S. M., and Ose, D. (2011). Predictors of health-related quality of life in patients at risk for cardiovascular disease in European primary care. *PLoS ONE, 6*(12), [e29334].
- Mahler, C., Rochon, J., Karstens, S., Szecsenyi, J., and Hermann, K. (2014). Internal consistency of the readiness for interprofessional learning scale in German health care students and professionals. *BMC Medical Education, 14*(1), [145].
- Moritz, S., Rochon, J., Völkel, S., Hilker, M., Hobbhahn, J., and Arlt, M. (2010). Determinants of cerebral oximetry in patients undergoing off-pump coronary artery bypass grafting: An observational study. *European Journal of Anaesthesiology, 27*(6), 542–549.
- Muth, C., Beyer, M., Fortin, M., Rochon, J., Oswald, F., Valderas, J. M., Harder, S., Glynn, L. G., Perera, R., Freitag, M., Kaspar, R., Gensichen, J., and van den Akker, M. (2014). Multimorbidity's research challenges and priorities from a clinical perspective: the case of 'Mr Curran'. *European Journal of General Practice, 20*(2), 139–147.
- Muth, C., Harder, S., Uhlmann, L., Rochon, J., Fullerton, B., Güthlin, C., Erler, A., Beyer, M., van den Akker, M., Perera, R., Knottnerus, A., Valderas, J. M., Gerlach, F. M., and Haefeli, W. E. (2016). Pilot study to test the feasibility of a trial design and complex intervention on PRIoritising MUltimedication in Multimorbidity in general practices (PRIMUMpilot). *BMJ Open, 6*, [e011613].
- Muth, C., Uhlmann, L., Haefeli, W. E., Rochon, J., van den Akker, M., Perera, R., Güthlin, C., Beyer, M., Oswald, F., Valderas, J. M., Knottnerus, J. A., Gerlach, F. M., and Harder, S. (2018). Effectiveness of a complex intervention on Prioritising Multimedication in Multimorbidity (PRIMUM) in primary care: results of a pragmatic cluster randomised controlled trial. *BMJ Open, 8*, [e017740].
- Muth, C., van den Akker, M., Blom, J. W., Mallen, C.D., Rochon, J., Schellevis, F. G., Becker, A., Beyer, M., Gensichen, J., Kirchner, H., Perera, R., Prados-Torres, A., Scherer, M., Thiem, U., van

den Bussche, H., and Glasziou, P. P. (2014). The Ariadne principles: how to handle multimorbidity in primary care consultations. *BMC Medicine, 12,* [223].

- Noest, S., Ludt, S., Klingenberg, A., Glassen, K., Heiss, F., Ose, D., Rochon, J., Bozorgmehr, K., Wensing, M., and Szecsenyi, J. (2014). Involving patients in detecting quality gaps in a fragmented healthcare system: development of a questionnaire for Patients' Experiences Across Health Care Sectors (PEACS). *International Journal for Quality in Health Care, 26*(3), 240–249.
- Nonnenmacher, C., Dalpke, A., Rochon, J., Flores-de-Jacoby, L., Mutters, R., and Heeg, K. (2005). Real-time polymerase chain reaction for detection and quantification of bacteria in periodontal patients. *Journal of Periodontology, 76*(9), 1542–1549.
- Oberacher-Velten, I., Prasser, C., Rochon, J., Ittner, K. P., Helbig, H., and Lorenz, B. (2011). The effects of midazolam on intraocular pressure in children during examination under sedation. *British Journal of Ophthalmology, 95*(8), 1102–1105.
- Ose, D., Rochon, J., Campbell, S. M, Wensing, M., Freund, T., van Lieshout, J., Längst, G., Szecsenyi, J., and Ludt, S. (2014). Health-related quality of life and risk factor control: the importance of educational level in prevention of cardiovascular diseases. *European Journal of Public Health, 24*(4), 679–684.
- Ose, D., Rochon, J., Campbell, S. M., Wensing, M., van Lieshout, J., Uhlmann, L., Freund, T., Szecsenyi, J., and Ludt, S. (2012). Secondary prevention in patients with coronary heart diseases: What factors are associated with health status in usual primary care? *PLoS ONE, 7,* [e51726].
- Rochon, J., Gondan, M., and Kieser, M. (2012). To test or not to test: Preliminary assessment of normality when comparing two independent samples. *BMC Medical Research Methodology, 12*, [81].
- Rochon, J., and Kieser, M. (2011). A closer look at the effect of preliminary goodness-of-fit testing for normality for the one-sample t-test. *British Journal of Mathematical and Statistical Psychology, 64*(3), 410–426.
- Roehl, I., Beyer, M., Gondan, M., Rochon, J., Bauer, E., Kaufmann-Kolle, P., and Gerlach, F. M. (2013). Guideline compliance in the treatment of patients with chronic heart failure through family doctor-centred health care: results of an evaluation study. *Zeitschrift fur Evidenz, Fortbildung und Qualität im Gesundheitswesen, 107*(6), 394–402.
- Sassen, A., Rochon, J., Wild, P., Hartmann, A., Hofstädter, F., Schwarz, S., and Brockhoff, G. (2008). Cytogenetic analysis of HER1/EGFR, HER2, HER3 and HER4 in 278 breast cancer patients. *Breast Cancer Research, 10*(1), [R2].
- Schnitzbauer, A. A., Scherer, M. N., Rochon, J., Sothmann, J., Farkas, S. A., Loss, M., Geissler, E. K., Obed, A., and Schlitt, H. J. (2010). Study protocol: a pilot study to determine the safety and efficacy of induction-therapy, de novo MPA and delayed mTOR-inhibition in liver transplant recipients with impaired renal function. PATRON-study. *BMC Nephrology, 11,* [24].
- Schnitzbauer, A. A., Zuelke, C., Graeb, C., Rochon, J., Bilbao, I., Burra, P., de Jong, K. P., Duvoux, C., Kneteman, N. M., Adam, R., Bechstein, W.O., Becker, T., Beckebaum, S., Chazouillières, O., Cillo, U., Colledan, M., Fändrich, F., Gugenheim, J., Hauss, J. P., Heise, M., Hidalgo, E., Jamsieson, N., Königsrainer, A., Lamby, P., Lerut, J.P., Mäkisalo, H., Margreiter, R., Mazzaferro, V., Mutzbauer, I., Otto, G., Pageuax, G. P., Pinna, A., Pirenne, J., Rizell, M., Rossi, G., Rostaing, L., Roy, A., Turrion, V. S., Schmidt, J., Troisi, R. I., van Hoek, B., Valente, U ., Wolf, P., Wolters, H., Mirza, D. F., Scholz, T., Steininger, R., Soderdahl, G., Strasser, S. I., Jauch, K. W., Neuhaus, P.,

Schlitt, H. J., and Geissler, E.K. (2010). A prospective randomised, open-labeled, trial comparing sirolimus-containing versus mTOR-inhibitor-free immunosuppression in patients undergoing liver transplantation for hepatocellular carcinoma. *BMC Cancer, 10,* [190].

- Spitalewsky, K., Rochon, J., Ganzinger, M., and Knaup, P. (2013). Potential and requirements of IT for ambient assisted living technologies: Results of a Delphi Study. *Methods of Information in Medicine, 52*(3), 231–238.
- Weinhold, J., Hunger, C., Bornhäuser, A., Link, L., Rochon, J., Wild, B., and Schweitzer, J. (2013). Family constellation seminars improve psychological functioning in a general population sample: results of a randomized controlled trial. *Journal of Counseling Psychology, 60*(4), 601–609.

Acknowledgments

I would like to express my deepest gratitude and appreciation to all the people who have played a significant role in the completion of my thesis.

First and foremost, I am truly humbled and honored to have had the opportunity to work under the guidance of my esteemed supervisor, Prof. Meinhard Kieser, on the topic of this dissertation during my time at the Institute of Medical Biometry in Heidelberg. I had the distinct privilege of not only acquiring extensive knowledge about various methodologies but also cultivating valuable leadership skills, all within an enriching environment. I cannot express enough how deeply thankful I am for the continuous support, encouragement, and inspiration that Prof. Kieser has provided throughout my journey. His belief in me, especially during the final stretch, has been truly instrumental in my accomplishment. Thank you, Prof. Kieser, for your exceptional mentorship and for playing such a pivotal role in shaping my growth and development.

I am most grateful to Prof. Andreas du Bois and the Arbeitsgemeinschaft Gynaekologische Onkologie (AGO) Study Group OVAR for the wonderful support of my work. The trust and generosity that the AGO-OVAR showed to me enabled me to become part of something bigger than me, helping many women with ovarian cancer, the 'silent killer', to get better treatment and survive. Andreas, you and your AGO-OVAR colleagues, opened so many doors to me when giving me access to your world, involving me in clinical trials, inviting me to contribute to expert groups, and allowing me to use the QS-OVAR data for my thesis. I will never forget it.

I also want to thank Prof. Theis Lange for the collaboration and his methodological advice which contributed to the successful development and implementation of new methodology for this dissertation. I will always remember our first catch up which was the beginning of a great professional collaboration and a true friendship.

Last but not least, I would like to extend my heartfelt thanks to my husband Matthias and my friends for their constant understanding and patience. Their unwavering belief in me has been a driving force behind my perseverance.

Eidesstattliche Versicherung

1. Bei der eingereichten Dissertation zu dem Thema “Impact of Institutional Research Activity on Quality of Care and Patient Outcomes in Ovarian Cancer” handelt es sich um meine eigenständig erbrachte Leistung.
2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.
3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.
4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.
5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Ort und Datum Unterschrift